

From: Swope, Sheridan
Sent: Thursday, June 09, 2005 5:56 PM
To: STIC-Biotech/ChemLib
Subject: 10/649,273

For 10/649,273, pls search and interference search:

SID 2: full-length against the NT data bases
148-414 against the NT data bases
176-414 against the NT data bases

Sheridan Swope, Ph.D.
Patent Examiner, AU 1652
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E02C70 Remsen Bld (Mailbox)

Paul Schulwitz

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: 6/20 _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2005, 15:19:45 ; Search time 4951.35 Seconds

(without alignments)
4051.513 Million cell updates/sec

Title: US-10-649-273-2

Perfect score: 2125
Sequence: 1 MLILTKTAGVFPKRSKRKV.....DISKEVGEASIKVQLQMEI 414

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=x1p
-Q/cgm2.1/USPTO.spool.p/US10649273/runat.15062005.111416.6030/app.query.fasta_1.1429
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPT=0 -LOPEXT=0
-INITs-bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10649273 @CGM 1.1 7742 @runat.15062005.111416.6030 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:
1: gb ba:
2: gb htg:
3: gb ln:
4: gb om:
5: gb ov:
6: gb pat:
7: gb ph:
8: gb pl:
9: gb pr:
10: gb ro:
11: gb ste:
12: gb sy:
13: gb un:
14: gb vi:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2125	100.0	1908	9	BC011904 Homo sapi
2	2125	100.0	2197	6	AR428803 Sequence
3	2090.5	98.4	1387	6	AR428808 Sequence
4	2090.5	98.4	1387	9	HSA295148 Homo sapi

5	2088	98.3	1245	6	AX664697 Sequence
6	2088	98.3	1820	6	AX664695 Sequence
7	1944	91.5	2208	6	AX713716 Sequence
8	1944	91.5	2208	9	AK055441 Homo sapi
9	1835	86.4	1844	10	BC058172 Mus muscu
10	1747	82.2	1416	6	ARS41929 Sequence
11	1725	81.2	1546	10	BC078974 Sequence
12	1385	65.2	1526	6	AR428809 Sequence
13	1362	64.1	1522	5	BX934991 Gallus ga
14	1273	59.9	1017	10	BC038910 Mus muscu
15	1208	56.8	1558	5	BX930963 Gallus ga
16	1133.5	56.2	84115	9	AC013468 Homo sapi
17	1186.5	55.8	14364	6	AR428807 Sequence
18	995.5	46.8	249601	2	AC114153 Rattus no
19	995.5	46.8	308652	2	AC121478 Rattus no
20	970	45.6	256751	10	AC122925 Mus muscu
21	950.5	44.7	1109	10	BC051211 Mus muscu
22	938	44.1	860	5	BX930694 Gallus ga
23	827	38.9	875	6	C0721898 Sequence
24	751.5	35.4	121251	5	AL591593 zebrafish
25	714.5	33.6	1576	3	AY051882 Drosophill
26	714.5	33.6	1601	6	C0606432 Sequence
27	710	33.4	1474	3	AK113378 Clona int
28	709	33.4	117322	5	AL672217 zebrafish
29	700.5	33.0	1385	6	BD157102 Primer fo
30	700.5	33.0	1385	6	AX878239 Sequence
31	700.5	33.0	1385	6	AK027836 Homo sapi
32	687.5	32.4	3656	6	C0606431 Sequence
33	687.5	32.4	14679	2	AC018262 Drosophill
34	687.5	32.4	180263	3	AC010671 Drosophill
35	687.5	32.4	207432	3	AE003513 Drosophill
36	668.5	31.5	1443	8	AY024338 Arabidops
37	668.5	31.5	1474	8	AY117283 Arabidops
38	668.5	31.5	1567	8	AY063864 Arabidops
39	662.5	31.2	1557	8	AY084577 Arabidops
40	656	30.9	1672	8	AK070912 Oryza sat
41	629	29.6	571	6	BD154975 Primer fo
42	629	29.6	571	6	AX874913 Sequence
43	618	29.1	110000	1	AP065013 Continuation (9 of
44	577.5	27.2	260271	1	AE017258 Wolbachia
45	568.5	26.8	333800	1	SM591792 Sinorhizo

ALIGNMENTS

RESULT 1	BC011904	1908 bp	mRNA	linear	PRI 23-DEC-2003
LOCUS	BC011904				
DEFINITION	Homo sapiens O-6-ialoglycoprotein endopeptidase-like 1, mRNA (CDNA				
ACCESSION	BC011904.2	GI:40225818			
VERSION	BC011904				
KEYWORDS	MGC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1908)				
AUTHORS	Klausner R.D., Collins P.S., Wagner C.M., Schuler G.D., Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Altschul S.F., Zeeberg B., Buotow K.H., Schefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Nulya S.N., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettman M., Madan A., Young A.C., Shcherenko Y., Sanchez A., Whiting M., Madan A., Young A.C., Shcherenko Y., Boutard G.G., Blakesley R.W., Touchman J.W., Green E.D.,				

TITLE
JOURNAL
PUBLISHED
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 1908)
Strausberg, R.
Submitted (30-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT
Contact: MGC help desk
Email: cgabds-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, J., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stancir, S., Thomas, P.J., Touchman, J.W.,
Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: http://image.lnl.gov
Series: IRAL Plate: 28 Row: 1 Column: 22.
Location/Qualifiers

FEATURES
source

1. 1908
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:20293 IMAGE:4121450"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH_MGC_17"
/lab_host="DH10B-R"
/note="Vector: POT87"
1. 1908
/gene="OSGEP1"
/db_xref="LocusID:64172"
104. 1348
/gene="OSGEP1"
/product="OSGEP1 protein"
/protein_id="AAH11904.1"
/db_xref="GI:15080282"
/db_xref="LocusID:64172"
/translation="MLILTKAGVFPKSKRYEFLRNFPGTLFLKIVLGLET
SCDDTAADVDERGVGAHISOTVEALKTGGIVPPAOLREHRIORIVQELSLSS
GSPSLSAIATTIKPGLALSLGVLSFSLQIVGPKPPIPIHMEHAHLTRITNK
VERPVLVLSGHCILALVQVSDPRLILKSLDIAGDMLDVYARLSLTKRPGCT
WGGKAIHIAKGNRPFDIKPPLHAKGDSFTGLQVTDKILKKKEGIEKG
OILSSADIAATVQHTMAHILKTRTHRAILFCQRDILPNNNLVAVSGVANSFYLR
RALEILTNAOCTLCPPRLCTDNGIMIAMNGIERLRAGLILHIDIEGIRYKCPPL
GVDISKEVGEASIKVPLKMEI"
215. 1111
/gene="OSGEP1"
/note="Peptidase M22; Region: Glycoprotease family"
/db_xref="CD:pfam0814"

ORIGIN
Alignment Scores:

	Pred. No.:	Score:	Percent Similarity:	Best local Similarity:	Query Match:	DB:
	6.17e-171	2125.00	100.00%	100.00%	100.00%	9
						Gaps: 0
						Indels: 0
						Mismatches: 0
						Conservative: 414
						Matches: 414
						Length: 1908
						US-10-649-273-2 (1-414) x BC011904 (1-1908)
Qy	1	MetleuileuThrlystrAlaGlyValPhepheysProserlyArglyValTyr	20			
Db	104	ATGCTAATCTTGACTAGACGTCGAGAGTTTCTTTTAAACCATCAAGAAAGACTTAT	163			
Qy	21	GlupheuyrSerpheasnPhenipProGlyThrleupheuhislyleValleu	40			
Db	164	GAATTTTGAAGACTTTTAAATTTTCACTCGAAGACATTTCTTCAATTAATAGTATG	223			
Qy	41	GlytleuThrSerCyaspasPthAlaAlaValaValaPgluThrGlyAsnVal	60			
Db	224	GGAATTAAGAACTAGTTGATGATACAGACAGCTCTGCTGATGAACCTGGAATGTG	283			
Qy	61	leuGlyluAlaIleHisserGlnThrGluValHisleuysThrGlyGlyIleValPro	80			
Db	284	TTGGAGAGCAATACATCCCAACCTGAAGTTATTTAAACAGGCTGGAATGTCT	343			
Qy	81	ProAlaIaGlnGlnleuHisArgGluAsnIleGlnArgIleValGlnIuAlaLeuSer	100			
Db	344	CCAGACGCTCAGACGCTTCCAGAGAAATTTTCAAGAAATGACAAAGAGCTTTCT	403			
Qy	101	AlaSerGlyValSerProserAspLeuSerAlaIleAlaThrThrIleLysProGlyLeu	120			
Db	404	GCCAGTGAGCTCTCCAGAGTACCTCTCAGCAATGCAATACATTAACACAGACTT	463			
Qy	121	AlaIeuserleuGlyValGlyLeuSerPheSerleuGlnleuValGlyIleuLys	140			
Db	464	GCTTAAAGCCGAGAGGCTTATTCATTTAGCTTACAGCTGAGGACGTTAAAAAG	523			
Qy	141	ProPheIleProIleHisIleMetGluAlaHisAlaIleuThrIleArgLeuThrAsnLys	160			
Db	524	CCATTCAATCCCATTCATCATATGAGGCTCATGCACTTACTATTAGTTGACCAATAA	583			
Qy	161	ValGluPheProPheLeuValleuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuVal	180			
Db	584	GTAATATTTCTTTTATTTAGTTCTTTTATTTCTGAGGTCATCTGTGTGCAATAGT	643			
Qy	181	GlnGlyValSerAspPheLeuLeuGlyLysSerleuAspIleAlaProGlyAspMet	200			
Db	644	CAAGAGTTTCAGATTTTCTGCTTCTTGAAGCTTTGACATAGCACAGGTCACTAG	703			
Qy	201	LeuAspLysValAlaArgArgLeuSerleuIleLysHisProGluCysSerThrMetSer	220			
Db	704	CTTGACAAAGTGCGAGAGAGCTTTCTTTTAAATTAACCTCAAGAGTCTCCACATAGT	763			
Qy	221	GlyGlyLysAlaIleGlnHisleuAlaLysGlnGlyAsnArgPheHisPheAspIleLys	240			
Db	764	GCTGGAGAGCAATAGAAATTTGCGCAACAGAAATTAATTTGATTTGACATCAA	823			
Qy	241	ProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGlyLeuGlnHisValThr	260			
Db	824	CTCTCCCTGATCATGCTAAATAATTTGATTTCTTTTACCTGACCTCAACACCTTACT	883			
Qy	261	AspLysIleIleMetLysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly	280			
Db	884	GATTAATAATAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	943			
Qy	281	SerAlaIaAspIleAlaIaAlaThrValGlnHisThrMetAlaCysHisIleValLysArg	300			
Db	944	TGACGACGACATTTGCTGCTGACGACGACGACGACGACGACGACGACGACGACG	1003			
Qy	301	ThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaVal	320			
Db	1004	ACACATCGGCTTATTTCTTTTGTAGGAGAGAGAGCTTTTACCTCAAAATTAATGACAT	1063			

QY 321 LeuValAlaSerGlyValAlaSerAsnPheTyrIleArgAlaLeuGluIleLeu 340
 DB 1064 CTGGTTCATCTGTGGTGTGGCAAGTAACCTTCATATCCGACAGCTCTGGAAATTGTA 1123
 QY 341 ThrAsnAlaThrGlnCysThrIleuLeuCysProProAlaGluCysThrAspAsnGly 360
 DB 1124 ACAAGCGAACACAGTCACCTTGTGTGTCTCTCCAGACTATGACATGAATAAGGC 1183
 QY 361 IleMetIleAlaTyrAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAsp 380
 DB 1184 ATTAGATTCATCGAATGATGATTAAGAAAGACTAGCTGCTGGCATTTTATCATAC 1243
 QY 381 IleGluGlyIleArgTyrGluProLysCysProLeuGlyValAspIleSerLysGluVal 400
 DB 1244 ATGAAGGATCCGCTATGAACCAAAATGTCTCTTGAGATGACATATCAAAAGAGTT 1303
 QY 401 GlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 414
 DB 1304 GGAGAGCTTCATAAAGTACCAATTAATAATGAGATA 1345

 RESULT 2
 AR428803 AR428803 2197 bp DNA linear PAT 18-DEC-2003
 LOCUS Sequence 1 from patent US 6642041.
 DEFINITION AR428803
 ACCESSION AR428803
 VERSION AR428803.1 GI:40188589
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 2197)
 AUTHORS Chen, J., Feder, J.N., Nelson, T.C., Krystek, S.R. and Duclos, F.
 TITLE Polynucleotides encoding a novel metalloprotease, MP-1
 JOURNAL Patent: US 6642041-A 1 04-NOV-2003;
 FEATURES Location/Qualifiers
 source 1..2197
 /organism="Unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 7,37e-171 Length: 2197
 Score: 2125.00 Matches: 414
 Percent Similarity: 100.00% Conservatave: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-649-273-2 (1-414) x AR428803 (1-2197)
 QY 1 MetIleuIleLeuThrLysThrAlaGlyValPhePheLysProSerLysArgLysValTyr 20
 DB 231 ATGCTAATCTTGAAGTAAAGTACGACGAGAGTTTATTTTAAACATCAAAAAGAGTTTAT 290
 QY 21 GluPheLeuArgSerPheAsnPheHisProGlyThrIleuPheLysHisIleValLeu 40
 DB 291 GAATTTTAAAGAAATTTTAAATTTTCACTCGGAACACATATTTCTTCAAAAATGATATG 350
 QY 41 GlyIleGluThrSerCysAspAspThrAlaAlaValValAspGluThrGlyAsnVal 60
 DB 351 GGAATTAACCTAGTGTGATGATACAGCACTGCTGTGGTGAATGGAATGTG 410
 QY 61 LeuGlyGluAlaIleHisSerGlnThrGluValHisLeuLysThrGlyIleValPro 80
 DB 411 TTGGGAGACGAATACATTCCTCAAACTGAAGTTCATTTAAAAACAGTGGGATTTCTCT 470
 QY 81 ProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGluAlaLeuSer 100
 DB 471 CCAAGCACTCAACAGCTTCAACAGAAATATTCACCAAAATAGTACCAAGAGCTCTTTCT 530
 QY 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrTrpIleLysProGlyLeu 120
 DB 531 GCCAGTGAAGTCTCTCCAGATGACCTTCAGCAAAATGCAATACCAATAAACCAAGACTT 590

QY 121 AlaSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuLysLys 140
 DB 591 GCTTTAAGCCCTGGGAGGTGGCTTATCATTTGACTTACAGCTGTGAGCAGTAAAAAAG 650
 QY 141 ProPheIleProIleHisIleMetGlnAlaHisAlaLeuThrIleArgLeuThrAsnLys 160
 DB 651 CCATTCAATTCCTCATTCATATGAGGCTCATGACTTACTATTAAGGTGACCAATTA 710
 QY 161 ValGluPheProPheLeuValIleuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuVal 180
 DB 711 GTAGAAATTCCTTTTAAAGTTCTTTTGAATTTCTGAGGTCACTGCTGTGGCATTAAT 770
 QY 181 GlnGlyValSerAspPheLeuLeuLeuGlyLysSerLeuAspIleAlaIleProGlyAspMet 200
 DB 771 CAAGAGTTTCAGATTTTCTGCTCTTGGAAAGCTTTTGACATATGACACAGGTGACATG 830
 QY 201 LeuAspLysValAlaArgArgLeuSerIleLysHisIleProGluCysSerThrMetSer 220
 DB 831 CTGACAAAGGTGGCAAGAAAGACTTTCTTAATAAAACATCCAGAGTGTCTCCACCATGAGT 890
 QY 221 GlyGlyLysAlaIleGluHisLeuAlaLysGlnArgGlnArgPheHisPheAspIleLys 240
 DB 891 GGTGGAAAGCCATAGAACATTTGGCCAAACAGGAAATAGATTTTCATTTTGACATCA 950
 QY 241 ProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGlyLeuGlnHisValThr 260
 DB 951 CTCCCTTCATCATGCTAAATAATTTGATTTTCTTTTACGTGACCTTCAACAGCTTACT 1010
 QY 261 AspLysIleIleMetLysLysGluLysGluGluGlyIleGluLysGlyGlnIleLeuSer 280
 DB 1011 GATTAATAATTAATGAATAAAGGAAAGAGATATTTGAGAAAGGGCCAATCTGTCT 1070
 QY 281 SerAlaAlaAspIleAlaAlaThrValGlnHisIleThrMetAlaCysHisLeuValLysArg 300
 DB 1071 TCAGACAGACGATTTGCTGCCACAGTACAGACAAATGCGATGTCATCTGTGAAAGA 1130
 QY 301 ThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaVal 320
 DB 1131 ACAATCGGCTATTTCTGTTTGTGACAGAGACTTGTTCCTCAAAATTAATGACATG 1190
 QY 321 LeuValAlaSerGlyValAlaLaserAsnPheTyrIleArgAlaLeuGluIleLeu 340
 DB 1191 CTGGTTCATCTGGTGTGTGGCAAGTAACCTTCATATCCGACAGCTCTGGAATTTGTA 1250
 QY 341 ThrAsnAlaThrGlnCysThrIleuLeuCysProProAlaGluCysThrAspAsnGly 360
 DB 1251 ACAACGCAACACAGTGCATTTGTGTGTCTCTCCACAGCTATGACATGATATGTC 1310
 QY 361 IleMetIleAlaTyrAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAsp 380
 DB 1311 ATTAGATTTGATGAATGATTAAGTAAAGACTACCTGCTGTGGCATTTTATCATGAC 1370
 QY 381 IleGluGlyIleArgTyrGluProLysCysProLeuGlyValAspIleSerLysGluVal 400
 DB 1371 ATGAAGGCAATCCGCTATGAACCAAAATGTCTCTTGAGATGACATATCAAAAGAGT 1430
 QY 401 GlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 414
 DB 1431 GGAGAGCTTCATTAAGTACCAATTAATAATGAGATA 1472

 RESULT 3
 AR428808 AR428808 1387 bp DNA linear PAT 18-DEC-2003
 LOCUS Sequence 21 from patent US 6642041.
 DEFINITION AR428808
 ACCESSION AR428808
 VERSION AR428808.1 GI:40188594
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 1387)
 AUTHORS Chen, J., Feder, J.N., Nelson, T.C., Krystek, S.R. and Duclos, F.

TITLE Polynucleotides encoding a novel metalloprotease, MP-1
JOURNAL Patent: US 6642041-A 21 04-NOV-2003;
FEATURES Location/Qualifiers
source 1..1387
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.:	3,54e-168	Length:	1387
Score:	2090.50	Matches:	412
Percent Similarity:	93.85%	Conservative:	0
Best Local Similarity:	93.85%	Mismatches:	2
Query Match:	98.38%	Indels:	25
DB:	6	Gaps:	1

US-10-649-273-2 (1-414) x AR428808 (1-1387)

```

Oy 1 MetLeuIleLeuThrIleuThrAlaGlyValPhePheIysProSerIleSarGlyValIleTyr 20
Db 24 ATGCTATATCTGACTAAGACTGACGAGGTTTTTTTAAACATCAAAAGAAAGTTTAT 83
Oy 21 GluPheLeuArgSerPheAsnPheHisProGlyThrLeuPheLeuHisIleValLeu 40
Db 84 GAATTTTAAAGAGTTTAAATTTTCACTCTGAAACACTATTTCTTCAAAATAGTATTG 143
Oy 41 GlyIleGluThrSerCysAspAspThrAlaAlaValValAspGluThrGlyAsnVal 60
Db 144 GGAATTGAAACTGATGTGTGATGATGACGACGCTGTGTGTGATGAACTGGAAAGTGTG 203
Oy 61 LeuGlyValIleIleHisSerGlnThrGlyValHisIleuIysThrGlyIleValPro 80
Db 204 TTGGGAGAACCATATCACTTCCCAAGCTGAACTTATTAAGAGGTGGATTGTCTCT 263
Oy 81 ProAlaIleGlnGlnIleuHisSarGlyValAsnIleGlnArgIleValGlnIleValLeuSer 100
Db 264 CCAGGAGCTCAACAGCTTCAAGAGAAATATTCAAGAAATAGTACAAGAGCTTCTCT 323
Oy 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleIysProGlyLeu 120
Db 324 GCGAGTGGAGTCTCTCCAGTGCCTCTCAGCAATTCACATCCATMAAACCGAGACTT 383
Oy 121 AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnIleuValGlyIleuIysIys 140
Db 384 GCTTTAAGCTCGGAGAGGCGCTTATCTTATAGCTTACAGCTGGTGAACGCTTAAAG 443
Oy 141 ProPheIleProIleHisIleMetGlyAlaHisIleAlaLeuThrIleArgLeuThrAspIys 160
Db 444 CCATTCATTCCTCATTCATATGAGAGGCTCATGCACTTATAGTTGACCAATAAA 503
Oy 161 ValGluPheProPheLeuValIleuLeuIleSerGlyGlyHisIleCysLeuLeuAlaLeuVal 180
Db 504 GAGAAATTCCTTTTATGTTCTTTGATTTCTGAGAGTCACTGTCGTTGGCAATTAGTT 563
Oy 181 GlnGlyValSerAspPheLeuLeuLeuGlyIysSerIleuAspIleAlaProGlyIysMet 200
Db 564 CAGGAGTTTCAATTTTCTGCTTCTTGAAGAGCTTGGACATGACCAAGGAGCATG 623
Oy 201 LeuAspIleValAlaIleArgArgLeuSerLeuIleIysHisProGlyCysSerThrMetSer 220
Db 624 CTTGACAGGTGCGAAGAAACCTTCTTTTATTAAGAAACATCCAGAGTCTCCACATAGT 683
Oy 221 GlyGlyIysAlaIleGluHisIleuAlaIysGlnGlyAsnArgPheHisIleAspIleIys 240
Db 684 GGGGAGAAAGCCATAGAGCATTTGGCCAAACAAGAAATGATTTCAATTTTGAATCAAA 743
Oy 241 ProProLeuHisHisIleAlaIysAsnCysAspPheSerPheThrGlyLeuGlnHisIleValThr 260
Db 744 CTTCCCTTGATCATGCTMAAAATTTGATTTTCTTTTACCTGACTTCAACAGCTTACT 803
Oy 261 AspIysIleIleIleMetIysIysGlyIysGlyGlyIle----- 273
Db 804 GATTAATATATATATGAAAGGAAAGGAAAGGATATATTTCTTAATTAAGTAAAGTTGAA 863

```

```

Oy 274 -----GluIys 275
Db 864 CAGTAAATATTCCTGATGTTGCGCTAAATAATAGCTCATTTCTGACGAGTATGAGAAG 923
Oy 276 GlyGlnIleLeuSerSerAlaAlaAspIleAlaIleAlaThrValGlnHisThrMetAlaCys 295
Db 924 GGGCAATTCCTGCTTCAAGCAGACAGATGCTCCACATGACGACACAAATGCGCATGT 983
Oy 296 HisLeuValIysArgThrHisArgAlaIleLeuPheCysIysGlnArgAspLeuPro 315
Db 984 CATTTGTGAAAGAAACATCGGCTATTTCTGTTGTATAGCAGAGACTGTTACTT 1043
Oy 316 GlnAsnAsnAlaValIleuValAlaSerGlyGlyValAlaIleSerAsnPheThrIleArgArg 335
Db 1044 CAAATATATGACAGTACGTGTTGCACTGCTGTGTGTGCGAATTAATCTTATATCCGACGA 1103
Oy 336 AlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLeu 355
Db 1104 GCTCTGAAATTTTAAACAAAGCAACAGTGCACCTTGTGTGTCTCCAGACTA 1163
Oy 356 CysThrAspAsnGlyIleMetIleAlaIleArgAsnGlyIleGluArgLeuArgAlaGlyLeu 375
Db 1164 TGCACTATATATGCAATATATGATGCAATGATGATGATAAGACTACGCTGCTTG 1223
Oy 376 GlyIleLeuHisAspIleGluGlyIleArgTyrGluProIysCysProLeuGlyValAsp 395
Db 1224 GGCATTTTACATGACATAGAGGCACTCCGCTATGACCAAAATGTCTCTTGGAGTAC 1283
Oy 396 IleSerIysGlyValGlyValIleSerIleIysValProGlnLeuIysMetGluIle 414
Db 1284 ATATCAAAAGAGTTGGAGAAAGCTTCCATTAAGATACACATTAATAATGAGATA 1340

RESULT 4
HSA295148 1387 bp mRNA linear PRI 30-OCT-2000
LOCUS Homo sapiens mRNA for putative stialoglycoprotease type 2.
DEFINITION A295148
ACCESSION A295148
VERSION A295148.1 GI:11071726
KEYWORDS metalloprotease; stialoglycoprotease.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Carnifera; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Chen, J.M., Fortunato, M. and Barrett, A.J.
Cloning and sequencing of a second human putative
stialoglycoprotease homologue
unpublished
2 (bases 1 to 1387)
Chen, J.M.
Direct Submission
Submitted (27-OCT-2000) Chen J.M., MRC Molecular Enzymology
Laboratory, The Babraham Institute, Babraham, Cambridge, CB2 4AT,
UNITED KINGDOM
FEATURES
source Location/Qualifiers
1..1387
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="uterus"
24..1343
/function="metalloprotease of family M22"
/codon_start=1
/product="putative stialoglycoprotease type 2"
/protein_id="CAC14666.1"
/db_xref="GI:11071727"
/db_xref="GOA:Q9H4B0"
/db_xref="UniProt/TREMBL:Q9H4B0"
/citation="MLITRAGVFFKPSKRVFPLRSFNFHETFLAKIVGIER
SCDPTAAVVDENGVGEAIHSQTEVHLKTKGVPPAAQOIHENIORIYQALISAS
GVSPSDLSAATTKPGALISGVLSFSLQVQDKKPPRIHHMEHATLITLTK
VEPFLVLLISGHCILALVQGVSDPFLIGKSLDIATPDMLDKYARRSLIKHECST

```

ORIGIN

MSGKAIIEHLAQGNRFHFDIKPPLHAKNCDFSFTGLQHTVDKTIIMKEKERCIFLI
 SRVEQINIPGLCKI LAHFCRYEKQILSSADIAIVQITPMCHLVKTRHAIILFK
 QRDLEPNNNAVLCSGVSANFYIRRALIELTNATCTCLCPPLPCTDGMIMAWNG
 IRRRLGGLIDHIDIGIRYEPKCPGLGVIDISKEVGEASIKVQJLMEI"

Alignment Scores:

Pred. No.:	3,54e-168	Length:	1387
Score:	2090.50	Matches:	412
Percent Similarity:	93.85%	Conservative:	0
Best Local Similarity:	93.85%	Mismatches:	2
Query Match:	98.38%	Indels:	25
DB:	9	Gaps:	1

US-10-649-273-2 (1-414) x HSA295148 (1-1387)

QY 1 MetLeuIIeLeuThrlYsThrAlaGlyValPhePheLysProSerLysArgLysValTyr 20
 DB 24 ATGCTAATCTTTCAGTAACTGCGAGAGTTTTTTTAAACCAATCAAAAGAAAGTTTAT 83
 QY 21 GluPheLeuArgSerPheAsnPheHisProGlyThrlLeuPheLysIleValLeu 40
 DB 84 GAATTTTAAAGAGTTTAAATTTTCACTGAAACACTATTTCTTCAATAATAGTATG 143
 QY 41 GlyIleGluThrSerCysAspAspThrAlaAlaValAspGluThrGlyAsnVal 60
 DB 144 GGAATTTGAAACTAGTGTATGATACAGCAGCTGCTGTGATGAACTGAAATGCTG 203
 QY 61 LeuGlyGluAlaIleHisSerGlnThrGluValHisLeuLysThrGlyIleValPro 80
 DB 204 TTGGGAGAAACATACTCCCAACTGAAGTTTCATTTAAAAAACAGGTGGGATTTCTT 263
 QY 81 ProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGluAlaLeuSer 100
 DB 264 CCAGCGAGCTCAACAGCTTCCACAGAAATATTCAGAAATAGTACAGAGCTCTTCT 323
 QY 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrTrlIleLysProGlyLeu 120
 DB 324 GCCAGTGGAGTCTCTCCAGTGCCTCTCAGCAATTCACCACTACATTAACCCAGGACTT 383
 QY 121 AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuLys 140
 DB 384 GCTTTAAGCTGCGAGTGGCTTATATCTTAGCTTACAGCTGATAGGACGTTTAAAG 443
 QY 141 ProPheIleProIleHisMetGluAlaHisIleAlaLeuThrIleArgLeuThrAsnLys 160
 DB 444 CCATTCATTCCTCATTCATATGAGAGCTCATGCACTTACTATTAGGTTGACCAATAAA 503
 QY 161 ValGluPheProPheLeuValLeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuVal 180
 DB 504 GTAGAAATTCCTTTTATGTTCTTTGATTTCTGGAGGTCACTCTGTTGGCATTTAGTT 563
 QY 181 GlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMet 200
 DB 564 CAAAGAGTTTCAATTTTCTGCTTCTTGGAAAGCTTTTGACATACAGCAAGGTGACATG 623
 QY 201 LeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlyCysSerThrMetSer 220
 DB 624 CTTGACAGGTGCAAGAAACATTTCTTTAATAAATCATCCAGAGTCTCCACCATAGT 683
 QY 221 GlyGlyLysAlaIleGluHisLeuAlaLysGlnGlyAsnArgPheHisPheAspIleLys 240
 DB 684 GGTGGGAAAGCCATAGACATTTGGCCAAACAAAGAAATAGATTTCTTTTGAATCAAA 743
 QY 241 ProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGlyLeuGlnHisValThr 260
 DB 744 CTCCTCCCTGATATGCTAAATAATGTGATTTTCTTTTACGTGACCTCAACAGCTTACT 803
 QY 261 AspLysIleIleMetLysLysGluLysGluGlnGlyIle 273
 DB 804 GATTAATATATATGAAAAAGAAAAAGAGAGGTATATTTCTTAATTAGTAAAGTTGAA 863
 QY 274 -----GluLys 275

DB 864 CAGATTAATATTTCTGATGTTGCTTAAATAATACCTGCTCATTTCTGACAGTATGAGAG 923
 QY 276 GlyGlnIleLeuSerSerAlaAlaAspIleAlaIleAlaThrValGlnHisIleThrMetAlaCys 295
 DB 924 GGGCAATTCCTGCTTTCAGCAGCAGACATTCCTCCACAGTACAGACATATGGCATGT 983
 QY 296 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 315
 DB 984 CATCTTGAAAAAACAACATCGGGCTATCTGTTTGTAGACAGAGACTTGTACTT 1043
 QY 316 GlnAsnAsnAlaValIleLeuValAlaSerGlyGlyValAlaAspAsnPheTyrIleArgArg 335
 DB 1044 CAATAATATGCAATGCTGTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1103
 QY 336 AlaLeuGlnIleLeuThrAsnAlaIleThrGlnCysThrLeuLeuCysProProArgLeu 355
 DB 1104 GCTCTGGAATTTTAAACAAACGACACACAGTGCATTTGTTGTCTCTCTCCACGACTA 1163
 QY 356 CysThrAspAsnGlyIleMetIleAlaIleArgAsnGlyIleGluArgLeuArgAlaGlyLeu 375
 DB 1164 TGCACTGATATGCACTTATGATTCATGATGATGATGATGATGATGATGATGATGATGATG 1223
 QY 376 GlyIleLeuHisAspIleGluGlyIleArgTyrGluProLysCysProLeuGlyValAsp 395
 DB 1224 GGCATTTTACATGACATPAGAGGATCGCTATGACCAAAATGTCTCTTGAGTAGAC 1283
 QY 396 IleSerLysGluValGlyGluAlaSerTrlIleValProGlnLeuLysMetGluIle 414
 DB 1284 ATATCAAAAGAAAGTTTGAGAGAGCTTCCATTAAGTACCAATTAATAATGAGATA 1340

RESULT 5
 AX664697 LOCUS 1245 bp DNA linear PAT 22-MAR-2003
 DEFINITION Sequence 6 from Patent WO02074960.
 AX664697
 VERSION AX664697.1 GI:29164457
 KEYWORDS

SOURCE
 ORGANISM
 Homo sapiens (human)

REFERENCE
 AUTHORS
 Leiby K.R., Kapeller-Libermann R. and Glucksmann M.
 TITLE
 Mammalia; Eutheria; Chordata; Craniata; Vertebrate; Euteleostomi;
 Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.

JOURNAL
 38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions
 of human proteins and uses thereof
 Patent: WO 02074960-A 6 26-SEP-2002;
 Millennium Pharmaceuticals, Inc. (US)

FEATURES
 source
 1.1245
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	5.03e-168	Length:	1245
Score:	2088.00	Matches:	407
Percent Similarity:	99.03%	Conservative:	3
Best Local Similarity:	98.31%	Mismatches:	4
Query Match:	98.26%	Indels:	0
DB:	6	Gaps:	0

US-10-649-273-2 (1-414) x AX664697 (1-1245)

QY 1 MetLeuIIeLeuThrlYsThrAlaGlyValPhePheLysProSerLysArgLysValTyr 20
 DB 1 ATGCTAATCTTTCAGTAACTGCGAGAGTTTTTTTAAACCAATCAAAAGAAAGTTTAT 60
 QY 21 GluPheLeuArgSerPheAsnPheHisProGlyThrlLeuPheLysIleValLeu 40
 DB 61 GAATTTTAAAGAGTTTAAATTTTCACTGGAACATATTTCTTCAATAATAGTATTC 120

QY 41 GYIIEGIuThSerCySaPaPThrAlaAlaValaValaSpGIuThrGIYAsnVal 60
 DB 121 GGAATTGAACTAGTGTGTGTATGATACAGAGCTGTGTGTGTATGAAAGTGC 180
 QY 61 LeuGIYAlaAlaIleHisSerGIuThrGIYAlaIleuLeuThrGIYIleValPro 80
 DB 181 TTGGGGAAAGCATTACCTTCCCAACTGAGATTCTTTTAAACAGGTGGATTGTCT 240
 QY 81 ProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValaGlnGluAlaLeuSer 100
 DB 241 CCAGAGGCTCAACAGCTTCCACAGAAATATTCACAGATATGACAGAGCTTTCT 300
 QY 101 AlaserGIYAlaSerProSerAspLeuSerAlaIleAlaThrThrIleYsProGIYLeu 120
 DB 301 GCCAGTGGAGTCTCTCCAGTGAACCTCTCAGCAATTCACATCAATTAACAGAGCTT 360
 QY 121 AlaLeuSerLeuGIYAlaGIYLeuSerPheSerLeuGlnLeuValaGIYGlnLeuYsLeu 140
 DB 361 GCTTTAAGCTGGAGAGGGCTTATCATATTTAGCTTACGCTGTGTAGACAGTTAAAG 420
 QY 141 ProPheIleProIleHisIleHisMetGluAlaAlaIleuThrIleArgLeuThrAsnYs 160
 DB 421 CCATTATTCCTCATTCATCATATGAGAGCTCATGCACTTACTTATGAGTTGACCAATTA 480
 QY 161 ValaGluPheProPheLeuValaLeuLeuIleSerGIYAlaIleCysLeuLeuAlaLeuVal 180
 DB 481 GTAGAAATTCCTTTTATGTTTCTTTGATTTCTGAGAGTCACTGTCTGTGGCATTAAGT 540
 QY 181 GlnGIYAlaSerAspPheLeuLeuLeuGIYLeuSerLeuAspIleAlaProGIYAspMet 200
 DB 541 CAAGAGGTTTCAGATTTTCTGCTTCTTGAAAGTCTTTGACATAGACACAGAGTGACATG 600
 QY 201 LeuAspYsValaAlaArgArgLeuSerIleIleYsHisProGIYCysSerThrMetSer 220
 DB 601 CTTCGACAGGTGCACAGACATTTCTTTATTAACATTCACAGTGTCTCCACCATGAGT 660
 QY 221 GIYGIYAlaAlaIleGlnIleLeuAlaYsGlnGIYAsnArgPheHisPheAspIleYs 240
 DB 661 GGTGGGAAAGCCATAGAACATTTGGCCAAACAGAAATGATTTTCACTTTTGACATCAA 720
 QY 241 ProProLeuHisIleAlaYsAsnCySaPaPheSerPheThrGIYLeuGlnIleValThr 260
 DB 721 CCTCCCTTCATCATCTATAAAATGTGATTTTCTTTTCTGACATTCACACGTACT 780
 QY 261 AspYsIleIleMetYsGIYLeuGIYAsnArgPheHisPheAspIleYs 280
 DB 781 GATTAATAATATGATAAAACGAGAAACAGAGAAAGTATTAAGAGGGGCAAAATCTGTCT 840
 QY 281 SerAlaAlaAspIleAlaAlaThrValaGlnHisThrMetAlaCysHisLeuValaYsArg 300
 DB 841 TCAGCAGCAGACATTCCTGCCACAGTACAGCAACAAATGGCATGTCTCTTGGAAGAA 900
 QY 301 ThrHisArgAlaIleLeuPheCysYsGIYAsnArgPheLeuProGlnAsnAlaVala 320
 DB 901 ACACATCGGGCTATTCGTTTCTTAAGCAGAGAGACTGTGTACTCAAAATATATGCAGTA 960
 QY 321 LeuValaAlaSerGIYAlaAlaSerAsnPheYsIleArgArgAlaLeuGlnIleLeu 340
 DB 961 CTGGTTCATCTGT 1020
 QY 341 ThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeuCysThrAspAsnGIY 360
 DB 1021 ACMAAGCAGACACAGTGCCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
 QY 361 IleMetIleAlaIleArgAsnGIYIleGluArgLeuArgAlaGIYLeuGIYIleLeuHisAsp 380
 DB 1081 ATTATGATTCAGATGATGTGATTAAGACATGCTGTGTGTGTGTGTGTGTGTGTGTGT 1140
 QY 381 IleGluGIYIleArgArgLeuProYsCysProLeuGIYAlaIlePheSerGIYVala 400
 DB 1141 ATAGAAAGGATCCGCTATGAAACCAAAATGTCTCTTGAGATAGACATATCAAAAGAGTT 1200
 QY 401 GIYGIYAlaSerIleYsValaProGlnLeuYsMetGlnIle 414

DB 1201 GGAAGAGCTTCATAAAAGTACCACAATTAATAATGAGATA 1242
 RESULT 6
 LOCUS AX664695 1820 bp DNA linear PAT 22-MAR-2003
 DEFINITION Sequence 4 from Patent WO02074960.
 VERSION AX664695
 KEYWORDS AX664695.1 GI:29164455
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Leiby, K.R., Kapeller-Liberman, R. and Glucksmann, M.
 TITLE 38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions
 JOURNAL of human proteins and uses thereof
 Patent: WO 02074960-A 4 26-SEP-2002;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source location/Qualifiers
 1..1820
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 146..1390
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD80044.1"
 /db_xref="GI:29164455"
 /translation="MLITRTYAVFPKRSRKYEFILRSFNHPTLFLKIVGIET
 SCDTAAAVVDENGVGEAIHSQTEVHLKTGGIVPPAAQOHLRENIQRIVOEALNS
 GSPSDLSAATITIKPGLASLGSFSLQVGLKPEPLPIHREAHALITLTK
 VERPPLVTLISGCHLLAVGVSDPLILKSLIDAPGMDLKYARPLSLIKPEGST
 MSGKAIENLAKGNRHPDIKPLHAKKCDSPFTGLQVTTDKNNRQEBIEKG
 QILSSADIAATVQHTWACHLVKTRTHALILFCQQRDLIPNNAVLVASGVASPFYLR
 RALEIILTNATQCTLCPPRLCTDNGIMIMNGIERLRAGILGILHDIIGIRYBKPL
 GVDISKVGSASIKVPLKMEI"
 ORIGIN
 Alignment Scores:
 Pred. No.: 8,14e-168 Length: 1820
 Score: 2088.00 Matches: 407
 Percent Similarity: 99.03% Conservative: 3
 Best Local Similarity: 98.31% Mismatches: 4
 Query Match: 98.26% Indels: 0
 DB: Gaps: 6
 US-10-649-273-2 (1-414) x AX664695 (1-1820)
 QY 1 MetLeuIleLeuThrIysThrAlaGIYAlaPhePheYsProSerIysArgYsValaYr 20
 DB 146 ATGTAATCTTGAAGTAAAGTGCAGAGATTTTAAACCATCAAAAGGAAAGTTAT 205
 QY 21 GluPheLeuArgSerPheAsnPheHisProGIYThrPhePheLeuHisYsIleValLeu 40
 DB 206 GAATTTTAAAGAACTTTAATTTTCACTCCGGAACATATTTCTCTCAATAAAATGATTTG 265
 QY 41 GIYIIEGIuThSerCySaPaPThrAlaAlaValaValaSpGIuThrGIYAsnVal 60
 DB 266 GGAATTAACACTAGT 325
 QY 61 LeuGIYAlaAlaIleHisSerGIuThrGIYAlaIleuLeuThrGIYIleValPro 80
 DB 326 TTGGGAAACCAAAACATTTCCCAACTGAAGTTCAATTAATAACAGGTGGATGTCTCT 385
 QY 81 ProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValaGlnGluAlaLeuSer 100
 DB 386 CCAAGAGCTCAACAGCTTCCACAGAAATATTCACAGATATGACAGAGCTTTCTTCT 445
 QY 101 AlaserGIYAlaSerProSerAspLeuSerAlaIleAlaThrThrIleYsProGIYLeu 120
 DB 446 GCCAGTGGAGTCTCTCCAGTGAACCTCTCAGCAATTCACATCAATAAAACAGAGCTT 505

QY 121 AlaleuSerleuGlyValGlyLeuSerPheSerleuGlnleuValGlyGlnleuValys 140
Db 506 GCTTTAAGCTGGAGAGTGGGCTTATCTTACCTTACAGCTGATGAGCACTTAAAAAG 565
QY 141 ProPheIleProIleHisIleMetGlnAlaHisAlaLeuThrIleArgLeuThrAsnlys 160
Db 566 CCATTCATTCCTCCATTCATCATATGAGAGCTCATCTACTATTAAGGTACCAATTA 625
QY 161 ValGlnPheProPheLeuValleuLeuIleSerGlyGlyHisCysleuLeuAlaVal 180
Db 626 GTGAATTCCTCTTTTATGTTCTTTGATTTCTGAGGTCACTGCTGTTGCACTTACT 685
QY 181 GlnGlyValSerAspPheLeuLeuGlyLysSerleuAspIleAlaProGlyAspMet 200
Db 686 CAAGAGCTTCAGATTTTCTGCTCTTGGAAGCTTTGGACATACAGCAGGTGACATG 745
QY 201 LeuAspLysValAlaArgArgLeuSerleuIleLysHisProGluCysSerThreSer 220
Db 746 CTTGACAAAGGTGGCAAGAGCTTTCTTAATAAACAATCCAGAGTCTCCACATGAGT 805
QY 221 GlyGlyLysAlaIleGlnHisleuAlaLysGlnGlyAsnArgPheHisPheAspIleLys 240
Db 806 GTGGGAAAGCCATGAGCAATTTGGCCAAACAGAGAAATGATTTGATTTGACATCAA 865
QY 241 ProProleuHisHisAlaLysAsnCysAspPheSerPheThrGlyLeuGlnHisValThr 260
Db 866 CCTCCCTTCATCTGCTGATTAATAATTTGATTTTCTTTACTGAGCTTCAACACTTACT 925
QY 261 AspLysIleIleMetLysGlyLysGlnGlyGlnGlyIleGlnLysGlyGlnIleLeuSer 280
Db 926 GATTAATAATATGAAACAGAGAAACAGAGAGATATGAGAGGGCAATCCTGCT 985
QY 281 SerAlaAlaAspIleAlaIleThrValGlnHisThrMetAlaCysHisleuValLysArg 300
Db 986 TCAGACAGACATTCCTGCGCACAGTACAGACACAATGAGCATGCTCTTGAAAAAGA 1045
QY 301 ThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuProGlnAsnAsnAlaVal 320
Db 1046 ACAATCGGGCTATCTGTTGTTTGAAGCAGAGACTGTTACTCCAAATTAATGCAATA 1105
QY 321 LeuValAlaSerGlyValAlaLysAsnPheThrIleArgArgAlaLeuGlnIleLeu 340
Db 1106 CTGCTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1165
QY 341 ThrAsnAlaThrGlnCysThrLeuLeuCysAspProProArgLeuCysThrAspAsnGly 360
Db 1166 ACAAGCAACACAGTGCATTTGTTGTTGCTCTCCACAGCTATGACATGATATGCTC 1225
QY 361 IleMetIleAlaTrpAsnGlyIleGlnArgLeuArgAlaGlyLeuGlyIleLeuHisAsp 380
Db 1226 ATATATGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1285
QY 381 IleGlnGlyIleArgGlyTrpGluProLysCysProLeuGlyValAspIleSerLysGlnVal 400
Db 1286 ATAAAGGCACTCCCTATGAACCAAAATGCTCTTGGAAGTACATATCAAAAGAGTT 1345
QY 401 GlyGlnAlaSerIleLysValProGlnleuLysMetGlnIle 414
Db 1346 GGAAGACCTTCATTAAGTACCAATTAATAAATGAGATTA 1387

RESULT 7
AX713716 2208 bp DNA linear PAT 15-APR-2003
LOCUS Sequence 400 from Patent EP1293569.
DEFINITION AX713716
VERSION AX713716.1 GI:29888642
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
1

AUTHORS Isegai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and Masuko, Y.

TITLE Full-length cDNA

JOURNAL Patent: EP 1293569-A 400 19-MAR-2003; Helix Research Institute (JP); Research Association for Biotechnology (JP)

FEATURES
source location/Qualifiers
1..2208
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1,83e-155 Length: 2208
Score: 1944.00 Matches: 386
Percent Similarity: 93.24% Conservative: 0
Best Local Similarity: 93.24% Mismatches: 4
Query Match: 91.48% Indels: 24
DB: Gaps: 1

US-10-649-273-2 (1-414) x AX713716 (1-2208)

QY 1 MetleuIleleuThrLysThrAlaGlyValPhePheLysProSerLysArgLysValTyr 20
Db 344 ATGCTAATCTTGACATCAAGCTGACAGAGCTTTTAAACATCAAAAAAGAAAGTTTAT 403
QY 21 GluPheLeuArgSerPheAsnPheHisProGlyThrleuPheLeuHisLysIleValleu 40
Db 404 GAATTTTAAGAAGTTTAATTTTATTCATCTGGAACATATTTCTCATTAATAATGATTTG 463
QY 41 GlyIleGlnThrSerCysAspAspThrAlaAlaValAlaAspGlnThrGlyLysVal 60
Db 464 GGAATTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 523
QY 61 LeuGlyValAlaIleHisSerGlnThrGlnValHisleuLysThrGlyLysIleValPro 80
Db 524 TTGGAGAGAGCATATCATCTCCAACTGATGATTTTAAACAGAGTGGATTTGCTCT 583
QY 81 ProAlaAlaGlnGlnleuHisArgGlnLysnIleGlnArgIleValGlnGlnAlaLeuSer 100
Db 584 CCAGAGCTCAACAGCTTCAACAGAGAAATATTAACGAAATGATCAAGAGCTCTTCT 643
QY 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleLysProGlyLeu 120
Db 644 GCCAGTGAAGTCTCTCAAGTGAATCTTCAGCAATTTGCAATCAATTAACCAAGACTT 703
QY 121 AlaleuSerleuGlyValGlyLeuSerPheSerleuGlnleuValGlyGlnleuValys 140
Db 704 GCTTTAAGCTGGAGAGTGGCTTATCTTATGCTTACAGCTGATGAGAGCTTAAAAAG 763
QY 141 ProPheIleProIleHisIleMetGlnAlaHisAlaLeuThrIleArgLeuThrAsnlys 160
Db 764 CCATTCATTCCTCCATTCATCATATGAGAGCTCATCTTAATTAAGTGAACATAA 823
QY 161 ValGlnPheProPheLeuValleuLeuIleSerGlyGlyHisCysleuLeuAlaVal 180
Db 824 GTAGAAATTCCTTTTATGTTCTTTGATTTCTGAGAGTCACTGCTGTTGCAATTAAGT 883
QY 181 GlnGlyValSerAspPheLeuLeuGlyLysSerleuAspIleAlaProGlyAspMet 200
Db 884 CAAGAGGTTTCAGATTTTCTGCTCTTGGAAGTCTTGGAATGACACAGGTGACATG 943
QY 201 LeuAspLysValAlaArgArgLeuSerleuIleLysHisProGlnCysSerThreSer 220
Db 944 CTTGACAAAGGTGGCAAGAGCTTCTTAATAAACAATCCAGAGTCTCCACATGACT 1003
QY 221 GlyGlyLysAlaIleGlnHisleuAlaLysGlnGlyAsnArgPheHisPheAspIleLys 240
Db 1004 GTGGGAAAGCCATGAGCACTTTGGCCAAACAGAGAAATGATTTGATTTGACATCAA 1063

Oy		241	ProbleumSH.SAlalyasncYsaPheSerPhetnGlyLeuGl.NhiValThr	260
Db		1064	CCTCCCTTGCACTCATGCCTAAAAATTGTATTTTCTTTTA CTGGACCTTCACAACCGTACT	1122
Oy		261	AaplysIlellewelysglwlusglucugilyileguylvsgylemleuseR	280
Db		1124	GATAAATAATAAATGAAGAAAAGAAGAAAGATTA TGAGAAAGGCCAAATCCTGTCT	1183
Oy		281	SerLalaIsapiellelaIAleThrValelnHsthrMecAlaCyshieleVallyslarg	300
Db		1184	TGACCAGACGACATTGCTGCCAACGTACAGCACCAAAGGCACTGTCATCTGTGTAAAGGA	1243
Oy		301	ThrlsarGalalAleEupHeCySlySGlnArgAspLeuleuProGI.AsnasnAlayAI	320
Db		1244	ACAACITGGGGCTATTCTTTTTTTGTBAGACAGAGACTTTTRNCTCCAANAATATCAGCTA	1307
Oy		321	LcuvalAlasergiygilyvaialaseranpheTyrieargalaleuclunlleu	340
Db		1304	CTGGTTCATCTGGTGGTGTGGCAAGTAATCTTGTATCCGACAGACTCTGGAAATTTTAA	1363
Oy		341	ThrasnalatrrngInysThreuleuNCysPropoFroAgLeuCyeThTraspangly	360
Db		1364	ACAACCGCACACAGTGCATTTGTGTGTCTCTCCACAGCTATGACATGATATATGCG	1423
Oy		361	IlewtellaelatrpaengylielgulargeuaragalaglyleucluyilEleuhlsarp	380
Db		1424	ATTATGATTGGA-----	1435
Oy		381	IleglugliyIleaGrTYrgluPrloIycCysProleuGlyVaIAsplIeSerlysgIuvai	400
Db		1436	-----TGAGTCTCTCTTGSAAGRATGACATATCAAAAAGAGATT	1471
Oy		401	GlycyluaIserIlleyevailProGlnleulsNetGiulle	414
Db		1472	GGAGAGCTTCCATAAAGTAGCCACAATTAATAATGAGAGATA	1513
RESULT 8 AKO55441				
LOCUS				
DEFINITION			AKO55441 2208 bp mRNA linear PRI 30-JAN-2004	
ACCESSION			Homo sapiens cDNA FLJ30879 f18, clone FEBRA2004592, highly similar	
VERSION			to Homo sapiens mRNA for putative sialoglycoprotease type 2.	
KEYWORDS			AKO55441	
SOURCE			Oligo capping; fls (full insert sequence).	
ORGANISM			Homo sapiens (human)	
REFERENCE			Homo sapiens	
AUTHORS			Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 Oca,T., Suzuki,Y., Nishikawa,T., Otseuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Ohbayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ichii,S., Yamamoto,J., Saito.K., Kawai.Y., Isono.Y., Nakamura.Y., Nagabiri,K., Murakami,K., Yasuda.T., Iwayanagi,T., Wagasuma.M., Shiraori,A., Sudo.H., Hosoi,T., Kakui.Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoy.T., Furuya,T., Kikkawa,E., Omura.Y., Abe,K., Kamihara,K., Kasetsu,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi.T., Yamashta.H., Miracka.W., Fujimori.K., Tani.H., Kimeta,M., Watanabe.M., Hiiracha.S., Chiba.Y., Ichida.S., Ono.Y., Takiguchi,S., Watanabe.S., Yosida,M., Hocura,T., Kusano.J., Kanehori,K., Takahashi-Fuji,I.A., Harra,H., Tanabe,T., Nomura.Y., Togiya,S., Komai.F., Harra.R., Takeuchi,K., Arita,M., Imose,N., Muesashino,K., Yuuki.H., Oshima,A., Sasaki.N., Aotsuka.S., Yoshikawa.Y., Matsunuma.H., Ichihara.T., Shiohata.N., Sano.S., Moriya.S., Momiyama.H., Satoh.N., Takemi.S., Terashima.Y., Suzuki.O., Nakagawa,S., Senoo,A., Mizoguchi,H., Gocho.T., Shimizu,F., Wakebe,H., Hisighaki,H., Watanabe,T., Sugiyama,A., Takemoto.M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi.Y., Fujimori.Y., Komiyama.M., Tashiro.H., Tangami.A., Fujiwara.T., Ono.T., Yanada.K., Fujii.Y., Osaki,K., Hirao.M., Ohmor,i.Y., Kawabata,A., Hikiji.T., Kobatake,N., Itagaki,H., Ikema,Y., Okamoto,S., Oktiani,R., Kawakami,T., Noguchi.S., Itoh.T., Shigetake,K., Senba.T., Matsumura,K.,	

TITLE
JOURNAL PUBLISHED
REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES
source

CDS

ORIGIN

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:

DB:
US-10-649-273-2 (1-414) x AK055441 (1-2208)

Yy 1 MetleuileuThrllyThrAlaGlyValPhePhelyProserIlySarIglyValIyr 20
Db 344 ATGCTATCTTGTGCATAGACGTCAGAGAGTTTAAAAACCATCAAAAGAAGATTAT 403

Yy 21 GluPheLeuArgSerPheAsnPhenHisProCgLyThrLeuPheLeuHisIleValLeu 40

Nat. Genet. 36 (1), 40-45 (2004)
14702039

2 Nishi,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Sakto,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kaada,K., Nagatsuna,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuhio,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2208)
Isogai,T., Otsuki,T. and Sugiyama,T.
Direct Submissions
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:-81-438-52-13975, Fax:81-438-52-39886)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology (RAB); (supported by Japan Construction Association for Biotechnology (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

Location/Qualifiers
1..2208
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FEBRA2004592"
/issue_type="Brain"
/clone_lib="FEBRA2"
/dev_stage="fetus"
/note="Cloning vector: pME18SF13"
344..1438
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB70923.1"
/db_xref="GI:16550167"
/translation="MLITRTKNGCFPRFSKKRYEPFRSPFHFGTLFLKIYVIGIH
SCDPTAAAVVDENGLDELSHSOTVEHLKTGGIVPPAQOQLHEENTQIVQELSL
GVSPDSLAIATTIKPGALISLGVLSLDVOQLKPPIPHMHAHALTRLTNT
VFPEPVILISGHCILALVOGVSDPLILGKSLLDIAGDMIDKYARLPILIKPECS
WEGFKALEIAKOGNRFPDIKDPILHANONCDPEFTGLOHYTDKIIMKEKEBGEIE
OLSSAADIATVQHTMACHIYKRTTHRALIFLCCKRDILLPONNALVASGVASNFCL
RALRLITMAVTCTLCPPPLRCTDNGMIYA"

Length: 2208
Matches: 386
Conservative: 0
Mismatches: 4
Indels: 24
Gaps: 1

Db 404 GAATTTTAAAGATTATTTTCACTCCGAAACCTATTTCTTCAATAAATGATTTG 463
 Oy 41 G1Y1LeG1uThSerCySasPseThrAlaAlaValaValaSpG1uThrG1yAseVal 60
 Db 464 GGAATTGAAACTAGTTGTATGATACAGACGCTGCTGTGTGATGAATGAATG 523
 Oy 61 LeuG1yG1uAla1Leh1SerG1nThrG1uVal1H1e1Leu1yThrG1yG1y1Leu1Pro 80
 Db 524 TTGGGAGAAAGCAATACATTTCCCAACTGAAGTTCAATTAATAACAGGTGGATTGCTCT 583
 Oy 81 ProAla1a1aG1nG1uLeuH1sArG1uAen1LeG1nArG1Le1a1G1nG1uAlaLeuSer 100
 Db 584 CCAGACAGCTCAACAGCTTCAACAGAAATATTCACAGAAATAGTACAGAAAGCTTTCT 643
 Oy 101 AlaseG1yValSerProSerAspLeuSerAla1Lea1ThrThr1LeuYProG1yLeu 120
 Db 644 GCCAGAGGAGTCTCTCAAGTACCTCTGACCAATGCACTACATTAACCAAGACCTT 703
 Oy 121 AlaleuSerLeuG1yValG1yLeuSerPheSerLeuG1nLeuValG1yG1nLeuY1y1s 140
 Db 704 GCTTTAAGCTTGGAGTGGGCTTATCATTTAGCTTACAGCTGTAGGACAGTTAAAG 763
 Oy 141 ProPhe1LeuPro1Leh1Sh1eWetG1uAlaH1sAlaLeuThr1LeuArgLeuThrAsn1y1s 160
 Db 764 CCATTCATTCATTCATCATATGAGAGGCTCATGACCTTACTATTGAGTTGACCAATAAA 823
 Oy 161 ValG1uPheProPheLeuValLeuLeu1LeuSerG1yG1yH1sG1yLeuAlaLeuVal 180
 Db 824 GTAGAAATTCCTTTTATGATTTCTTTGATTTCTGAGAGTCACTGCTGTTGACATTAGTT 883
 Oy 181 G1nG1yValSerAspPheLeuLeuLeuG1yLeuSerLeuAsp1Lea1ProG1yAspMet 200
 Db 884 CAAAGAGTTTCAATTTTCTGCTTTCTTGGAAACTCTTTGACATAGACACAGGTCAGATG 943
 Oy 201 LeuAsp1yValAla1ArgArgLeuSerLeu1LeuYH1sAspProG1yCySerThrMetSer 220
 Db 944 CTGGAAGAAGTGGCAAGAAGCTTCTTTAATAAATCATCCAGAGTCTCCACCATAGT 1003
 Oy 221 G1yG1yLeuAla1LeG1uH1sLeuAla1yG1nG1yAsnArgPheH1sPheAsp1LeuY1s 240
 Db 1004 GGTGGGAAAGCCATACAACTTTGGCCAAACAAAGAAATGATTTGATTTGACATCAA 1063
 Oy 241 ProPro1LeuH1sH1a1yAsnCySasPseSerPheThrG1yLeuG1nH1sValThr 260
 Db 1064 CCTCCCTTCATCATCTATAAATTTGATTTCTTTTATCTGACATTCACACGTTACT 1123
 Oy 261 Asp1y1Le1LeuMet1y1sG1uY1sG1uG1uG1y1LeG1uY1sG1yG1n1LeuSer 280
 Db 1124 GATTAATAATTAATGAAAAAGAAAAAGAAAGGATTTAGAAAGGGGCAATCTCTCT 1183
 Oy 281 SerAla1a1aAsp1Lea1a1aThrValG1nH1sThrMetAlaCyH1sLeuVal1y1sArg 300
 Db 1184 TCAGACGACAGACTTCTGCTGCACAGTACACACACATGCACTCTTGTGAAGA 1243
 Oy 301 ThrH1sArgAla1Le1LeuPheCyS1yG1nArgAspLeuLeuProG1nAenAsnAlaVal 320
 Db 1244 ACAACATCGGGCTATCTGTTTCTTAAGACAGAGACTTGTTACTCTCAAAATAAGCAGTA 1303
 Oy 321 LeuValAlaSerG1yG1yVala1a1aSerAsnPheThr1LeuArgArgAlaLeuG1n1Leu 340
 Db 1304 CTGCTGTCATCTGCTGTGTGAGAAAGTCTGTAATCCGACAGAGCTCTGGAAATTTTA 1363
 Oy 341 ThrAsnAlaThrG1nCyThrLeuLeuCySProProProArgLeuCyThrAspAsnG1y 360
 Db 1364 ACMAAGCAACAGTCACTTTGTTGTGCTCTCTCCAGACTATGACATGATAAGGC 1423
 Oy 361 LeuMet1Lea1a1aTrpAsnG1y1LeG1uArgLeuArgAlaG1yLeuG1y1LeuH1sAsp 380
 Db 1424 ATATATGATGCA----- 1435
 Oy 381 TLeG1uG1y1LeuArgTyArg1uPPro1yCySProLeuG1yVala1a1aSer1yG1uVal 400
 Db 1436 -----TGATGCTCTTGTGAGTACATATCAAAAGAGTT 1471

Oy 401 G1yG1uAlaSer1LeuY1sValProG1nLeuY1sWetG1u1Le 414
 Db 1472 GGAGAAAGCTTCATTAAGTACCAATTAATAAATGAGATTA 1513

RESULT 9
 BC058172
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS

BC058172 1844 bp mRNA linear ROD 08-OCT-2003
 Mus musculus cDNA clone MGC:67870 IMAGE:5012054, complete cds.
 BC058172
 MGC:67870.1 GI:34849663
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1844)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Kodrzycki, S.,
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywicki, M.I., Skalek, U., Smailus, D.E.,
 Schnerch, A., Schein, J.B., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

2 (bases 1 to 1844)
 Strausberg, R.
 Direct Submission
 Submitted (15-SEP-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: gcgabs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILML)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooke, S.,
 Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, B., Kwong, P., Latic, P., Legaspi, R.,
 Maduro, Q.L., Mastello, C., Maestri, B., Mastrian, S.D., McCloskey, D.C.,
 McDowell, J., Pearson, R., Stantirip, S., Thomas, P.J., Touchman, J.W.,
 Tsurgoun, C., Vogt, J.L., Walker, M.A., Weherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILML at: <http://image.lln.gov>
 Series: IRAP Plate: 123 Row: d Column: 22
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 21312463.
 Location/Qualifiers

FEATURES


```

source
1..1844
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech 11"
/db_xref="taxon:10090"
/clone="MGC:67870 IMAGE:5012054"
/tissue_type="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clonal) outgrowth infected with the virus MTV."
/clone_1db="NCI_CGAP_Lu29"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
222..1466
/codon_start=1
/product="Unknown (protein for MGC:67870)"
/protein_id="AAH58172.1"
/db_xref="GI:34849664"
/translation="MLMRPFAAIPKPKSKVGFLLRFSVHPRITLSCHLVGIEF
SCDDTGAAYVDENGNVLGEALISOTQVHLKGTGIVPPAOLHRENIQVEETLSAS
RIPSPDSLAIATTKRGLALSLGGLSLSLOLVNRFKPPPIPHHMAHALTILTNK
VEPFLVLLISGGHCLALVGVSDPFLIGSLDIADPMIDKVARRLSLIKPECSST
MSGKALIEHLAKDGNRFPTINPMQAKNDFSEFTGIQHTDRLITHEKEBQIEKGK
QIISAAIDIAAVOHATACHTAKETHAILECKKNLISPANAIVAVSGVASNLKIRP
KALEIVANATQCTLLCPPEPRLCTNGMIANNGIERLRAGLGVGHVDEDIRYBKPCPL
GIDISREVAEAAIKVPRKXAL"
misc_feature
333..1397
/notes="ORF17; Region: Metal-dependent proteases with
possible chaperone activity [Posttranslational
modification, protein turnover, chaperones]"
/db_xref="CDD:COG0533"

ORIGIN

Alignment Scores:
Pred. No.:      2,71e-146      Length:      1844
Score:          1835.00        Matches:      352
Percent Similarity: 91.06%      Conservative: 25
Best Local Similarity: 85.02%      Mismatches:   37
Query Match:    86.35%         Gaps:         0
DB:             10             Indels:        0

US-10-649-273-2 (1-414) x BC058172 (1-1844)

QY      1 MetLeuIleLeuThrLysThrAlaGlyValPhePheLysProSerLysArgLysValTyr 20
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      222 ATGCTAATGTTAAGAGACAGACAGACGACTATTCCCAAGCCCCCAAGAGTAAAGTTTAT 281

QY      21 GluPheLeuArgSerPheAsnPheHisProGlyThrLeuPheLeuHisLysIleValLeu 40
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      282 GGATTTTAAAGATTAAAGATTAAAGTTTATCCCAAACTCTCTTGTCAATAACGGTCTCTG 341

QY      41 GlyIleGluThrSerCysAspAspThrLeuAlaAlaValValAspGluThrGlyAsnVal 60
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      342 GGAATTGAACACAGCTGTGATGACACAGACCGCTGTGTGATGAATAACCTGGAAATGTG 401

QY      61 LeuGlyGluAlaIleHisSerGlnThrGluValHisLeuLysThrGlyIleValPro 80
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      402 CTGGGGGAGACACTGCACCTCCCAACTCAGGTTCATCGAAACACAGGTGGGATTGTTCT 461

QY      81 ProAlaIaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGluAlaLeuSer 100
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      462 CCACTAGCTCAACAACCTTCACAGAAATAATATCAACGAATAATGTAAGAAACCTCTTCT 521

QY      101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleLysProGlyLeu 120
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      522 GCCAGTAGAATCACCCCAAGCGATCTCTCGGCAATTCGAACTACCATCAACCGGGACTG 581

QY      121 AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuLysLys 140
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      582 GCCCTAAGCTTAGAGGTGCGGCTTATCCTTAGCTTACAGCTAATGTAATACGGTTAAAG 641

QY      141 ProPheIleProIleHisIleMetGluAlaHisAlaLeuThrIleArgLeuThrAsnLys 160
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      642 CCGTTATCCCGAATTCATCATGAGGCTCACGACGACTAATTTAGGCTCACCAATTA 701

```

Qy	161	ValGlnpHeProPheLeuValIleLeuIIISerGlyGlyHisCysLeuLeuAlaIleVal	180
Db	702	GTMGAATTCCTTCTTCTAGTTCCTTTATATTTCTGGCGGTCACTGCTGTGTGGCATTTAGTC	761
Qy	181	GIAGlyValSerAspPheLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMet	200
Db	762	CAAGGTCTTCCGATTTCTCGTCTCTTGGGAAGCTTTTGACATATAGCGCGACATG	821
Qy	201	LeuAspLysValAlaArgArgLeuSerLeuIIleLysHisProGluCysSerThreSer	220
Db	822	CTTGACAAAGGTGGCAAGAACATTTCTTTATCAAAATCCAGATGCTTCAACATGATG	881
Qy	221	GIYGLYysValIleGluHisIleLeuAlaLysGlnGlyAsnArgPheHisSpheAspIIleLys	240
Db	882	GGTGGAAAGCTATAGAACATTTGGCCAAACGGAATATGATTCATTTTACTATCATAT	941
Qy	241	ProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGlyLeuGlnHisValThr	260
Db	942	CCACCTATGCAAGATGCTAAGAAATGGCATTTTCTTTACGGGACCTTCAACATATACT	1001
Qy	261	AspLysIleIIleMetLysLysGluLysGluGluGlyIIleGluLysGlyGlnIIleLeuSer	280
Db	1002	GATPAGCTAATTAACACAAAGAAAAAGAAAGAGCATTTGGAAGGGGCAATCTGTGCA	1061
Qy	281	SerAlaAlaAspIleAlaIleAlaThrValGIIHisThrMetAlaCysHisIleValLysArg	300
Db	1062	TCACTGCAACATGCTGCTGCGGTACAGATGCAACAGGTGTCACCTTGCAAAAAAGA	1121
Qy	301	ThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaVal	320
Db	1122	ACACATGCTGCTATTTCTGTTTGGCAACAGAAAAATTTGCTATCTCCAGCTTAACGCAGTA	1181
Qy	321	LeuValAlaSerGlyGlyValAlaIleSerAspPheYrIIleArgAlaIleGlnIIleLeu	340
Db	1182	TTAGTTTATCTGGAGGTGTTCGAAAGTAACTTGACATCCGAAAAAGCATGGAAATGTCTC	1241
Qy	341	ThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeuCysThrAspAsnGly	360
Db	1242	GCAATTCGAACGCGTGCACCTTCTGTGTCCCTCCAGACCTGTGCATGACATGCG	1301
Qy	361	IleMetIleAlaIleArgAsnGlyIIleGluArgLeuArgAlaGlyLeuGlyIIleLeuHisAsp	380
Db	1302	ATCATGATTCATGATGAATGGAATTAAGTAAGATTAAGTGTGGGCTTTTACATGAT	1361
Qy	381	IIleGluGlyIIleArgIYrGluProLysCysProLeuGlyValAspIIleSerLysGluVal	400
Db	1362	GTMGAACATCCCATATGGAACCAAAATGTCTCTTGGAATAAGACATATCCAGAAAGTT	1421
Qy	401	GlyGluAlaSerIleLysValProGlnLeuLysMetGluIle	414
Db	1422	GCAGAGCTGCCATTAATAAGTACCGCGATTAATAATGGCATT	1463
RESULT 10			
LOCUS	ARS41929	1416 bp	DNA
DEFINITION	Sequence 177 from patent US 6743619.	linear	PAT 08-OCT-2004
ACCESSION	ARS41929		
VERSION	ARS41929.1	GI:53934009	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1416)		
AUTHORS	Tang,Y.T., Zhou,P., Goodrich,R., Liu,C., Aeundi,V., Ren,F., Zhang,J., Zhao,Q.A., Yang,Y., Xue,A.J., Weinman,T., Wang,J.-R., Wang,D. and Dirmacac,R.T.		
TITLE	Nucleic acids and polypeptides		
JOURNAL	Patent: US 6743619-A 177 01-JUN-2004;		
FEATURES	location/Qualifiers		
SOURCE	1..1416		
	/organism="unknown"		
	/mol type="genomic DNA"		

ORIGIN

Alignment Scores:

Pred. No.:	5,9e-139	Length:	1416
Score:	1747.00	Matches:	340
Percent Similarity:	100.008	Conservative:	1
Best Local Similarity:	99.71%	Mismatches:	0
Query Match:	82.21%	Indels:	0
DB:	6	Gaps:	0

US-10-649-273-2 (1-414) x AR541929 (1-1416)

```

Oy 74 LysThrGlyGlyIleValProProAlaIleGlnIleuHisArgGluAsnIleGlnArg 93
Db 280 AGAAGAGGTGGATTGTTCTCCAGCAGCTCAACAGCTTCACAGAGAAATATTCACAGCA 339
Oy 94 ILeValGlnGluAlaLeuSerAlaSerGlyValSerProSerAspLeuSerAlaIleAla 113
Db 340 ATAGTACAGAGACCTCTTCTGCGCAGTGAGTCTCCAGTGAACCTCTCAGCAATTGCA 399
Oy 114 ThrThrIleLeuSerProGlyLeuAlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGln 133
Db 400 ACTACCATTAACACAGAGACTTGTTAAGCTGGAGTGGAGCTTATCATTTAGCTTACAG 459
Oy 134 ILeuValGlyGlnLeuValSerProPheIleProIleHisIleMetGluAlaHisAlaLeu 153
Db 460 CTGGTAGCAGCATTTAAAGAACCTTCATTCCTCATCATATGAGGCTCATGCACTT 519
Oy 154 ThrIleArgIleuThrAsnIleValGluPheProPheLeuValLeuLeuIleSerGly 173
Db 520 ACTATTAGTTGACCAATTAAGTGAATTTCTTTTATGATTTGATTTCTGAGAGT 579
Oy 174 HisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuLeuGlyValSerLeu 193
Db 580 CACTGCTGTGGCATTAAGTTCAGAGGATTCAGATTTTCCTTCTTGAAGACTTTTG 639
Oy 194 AspIleAlaProGlyAspMetLeuAspValAlaArgArgLeuSerLeuIleValHis 213
Db 640 GACATTCACCAAGTGCATGCTTGCACAGGTGGCAGAGACCTTCTTTAATAAATCAT 699
Oy 214 ProGluCysSerThrMetSerGlyGlyValAlaIleGluHisLeuAlaValGlnGlyAsn 233
Db 700 CCAAGAGTCCCAACCATGAGTGGTGGGAAAGCCATGAAACATTTGGCCAAACAGAAAT 759
Oy 234 ArgPheHisPheAspIleValSerProProLeuHisIleAlaValSerCysAspPheSerPhe 253
Db 760 AGATTTTCAATTTGACATCAAACTCCCTTGATCATGCTTAATAATTTGATTTTCTTTT 819
Oy 254 ThrGlyLeuGlnHisValThrAspValIleIleMetValSerGluValGluGlyIle 273
Db 820 ACTGGACTTAAACCGTTACTGATAATTAATTAATGAAGAAAGAGAAAGAGATAT 879
Oy 274 GluValSerGlnIleLeuSerSerAlaAlaAspIleAlaIleThrValGlnHisThrMet 293
Db 880 GAGAGAGGGCAATTCCTGCTTCCAGCAGCAGACTTGTGCAAGTACAGACACATG 939
Oy 294 AlaCysHisLeuValIleValArgThrHisArgAlaIleLeuPheCysValArgAspLeu 313
Db 940 GCATGTCACTTGTGAAGAAAGACATCGGGCTATTGTTTGTGAACAGAGACTTG 999
Oy 314 LeuProGlnAsnAsnAlaValLeuValAlaSerGlyValAlaAlaSerAsnPheTyrlle 333
Db 1000 TTACCTCAAAATATATGACATGCTGTTGATCTGCTGTCGCAAGTACCTTCAATATC 1059
Oy 334 ArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProPro 353
Db 1060 CGAGAGCTCTGGAAATTTTAAACAGCAACACAGTGCACCTTGTGTGTCTCTCC 1119
Oy 354 ArgLeuValThrAsnArgIleValMetIleAlaIleArgAsnGlyIleGluArgLeuArgAla 373
Db 1120 AGACTATGCACTGATTAATGCAATTAATGATTCATGCAATGATGATTAAGAAAGACTAGTCT 1179
Oy 374 GlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrgluProValCysProLeuGly 393

```

```

Db 1180 GACCTGGGCAATTTACATGACATGACATGACATCCGCTATGAACCAAAATGCTCTTGA 1239
Oy 394 ValAspIleSerIleGluValAlaGlyAlaSerIleValProGlnLeuValSerMetGlu 413
Db 1240 GTAGCATATATCAAAAGAAAGTTGAGAACCTTCATTAAGTACCAATTAATAATGAG 1299
Oy 414 ILe 414
Db 1300 ATA 1302

RESULT 11
BC078974 1546 bp mRNA linear ROD 03-AUG-2004
LOCUS Rattus norvegicus cDNA clone IMAGE:7111906, partial cds.
DEFINITION BC078974
ACCESSION BC078974
VERSION BC078974.1 GI:50925879
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 1546)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldi,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Umedin,T.B., Toshimiyu,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,R.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huliyil,S.M.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Hellon,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmitz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skaleja,U., Smalins,D.E.,
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1546)
DIRECTOR MGC Project.
TITLE Submitted (02-AUG-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Howard Jacobs
CDNA Library Preparation: Express Genomics
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNU)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmitz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNU, at: http://image.llnl.gov
Series: rxn_plate: 182 Row: 6 Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
FEATURES

```


source
1. 1546
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:7111906"
/tissue_type="Testis, rat (Brown Norway)"
/clone_id="NH_MGC_237"
/lab_host="DH10B"
/note="Vector: pExpress1"
144. .>1546
/codon_start=1
/product="Unknown (protein for IMAGE:7111906)"
/protein_id="AAH78974.1"
/db_xref="GI:50926880"
/translation="MMLSTKAGAPRRPRNVRGPIRRFNVPRALEHKLVIQIET
SCDDPAAVDENVLGEALHSOTRHILKTGIVPVPAQOLRENTORIVLEALSAS
GVSPDLASIAITTKRGALSLGVGSFSQVLOVFRKPIPIHHEAALITRLTHK
VGPFLVLISGHCLLALVQSVDFLLKSLDIAFGMDLVARLSLIKPECT
MSGKAIIEHLAKENRPFHTINPMONAKNDFSFTGLQHTVDLITHKEBEGIEG
QILSSADIAAAVQATACHLAKRTHRALIFCOOKMLSPANNVLVSGVASNLVYR
RALEIVANATQCTLLCPRLCTDNGIMIMNGIERLRAGILIHVEDIIRYBKAGS
IYSMLERLALTELDLSVSTHTVHSPILNSSRGANIQISMCSCTQTVYMRVRL
NILNKKKKKKKKKKK"

ORIGIN

Alignment Scores:

Pred. No.:	4.9e-137	Length:	1546
Score:	1725.00	Matches:	335
Percent Similarity:	91.00%	Conservative:	19
Best Local Similarity:	86.12%	Mismatches:	35
Query Match:	81.18%	Indels:	0
DB:	10	Gaps:	0

US-10-649-273-2 (1-414) x BC078974 (1-1546)

QY 1 MetLeuIleuThrsThraIagIyValPhePheIysProSerIysSarGlyValTyr 20
DB 144 ATGCTATGTTAGTAAGACAGAGAGCTATCCGACACCAAGAGTAAGTTCGT 203
QY 21 GluPheLeuArgSerPheLeuPheHisProGlyThrLeuPheLeuHisIleValLeu 40
DB 204 GGAATTATTAAGATTTTATTTCACTTCAAGACCTTTTCACTTCAAGACCTGTCGT 263
QY 41 GlyIleGluThrsCysAspAspThrAlaAlaValAlaAspGluThrGlyAsnVal 60
DB 264 GGAATTGAACCAAGCTGTGATGACACAGAGCTGCTGTGATGATGAACCTGGAAGTGTG 323
QY 61 LeuGlyValAlaIleHisSerGlnThrGlyValHisIleLeuThrsGlyIleValPro 80
DB 324 CTGGGAGAGCACTGCCTCCAGACGGAAGTCCATCTGAACAGGTGGGATTTCTCT 383
QY 81 ProAlaAlaGlnGlnLeuHisSarGlyAsnIleGlnArgIleValGlnGlnValAlaLeuSer 100
DB 384 CCAGTGTCTCAACAGCTTCAAGAGAAATTTCAAGCAATAGTAAAGAAAGCTTTCT 443
QY 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleIysProGlyLeu 120
DB 444 GCCAGTGGGCTCTCCCAAGCACTCTCAGCAATTCACATCAACCAAGGACGTG 503
QY 121 AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuIleIys 140
DB 504 GCCCTAAGCTGGGGGTGGCTTATCTTTAGCGCTACAGTAACTAGTTTAAAG 563
QY 141 ProPheIleProIleHisIleMetGlyAlaHisAlaLeuThrIleArgLeuThrAsnIys 160
DB 564 CCATTTATCCCATTCATCAATGAGAGCTCAAGCGCTGATTTAGTCAAGCAACCAAAA 623
QY 161 ValGluPheProPheLeuValLeuLeuIleSerGlyIleIysIleValLeuAlaVal 180
DB 624 GTGCGATTCCTTTTATTTGATTTCTGGAGGCCACTGCTGGCCCTTATGTT 683
QY 181 GlnGlyValSerAspPheLeuLeuGlyIleSerLeuAspIleAlaProGlyAspMet 200

DB 684 CAGAGTGTTCAGATTTTCTGCTCTCGGAGATCCCTGACATAGCCCAAGGCAATG 743
QY 201 LeuAspIysValAlaArgArgLeuSerLeuIleIysHisProGlyCysSerThrMetSer 220
DB 744 CTTCACAGAGGTGGCAAGAGCTTTCTTTATCAACATCCAGATGTTTACATATGAT 803
QY 221 GlyIleIysAlaIleGlnHisIleValIleValIleValIleValIleValIleVal 240
DB 804 GTGGGAAAGCTATAGAACTTTGGCCAAAGAAATAGATTCACCTTATCATCAT 863
QY 241 ProProLeuHisHisAlaIleValAsnCysAspPheSerPheThrGlyLeuGlnHisValThr 260
DB 864 CCACCCATGAGATGCTAAGAACTGATTTTCTTTAGGAGCTTCAACATGTCACC 923
QY 261 AspIysIleIleMetIleIysGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGly 280
DB 924 GATTAAGCTAATTAACACACAGAAAGAAAGAAAGAGCTTGAAGAGGGCAATCTGTCA 983
QY 281 SerAlaAlaAspIleAlaIleThrValGlnHisThrMetAlaCysHisIleValIysArg 300
DB 984 TCAGCCGACAGCATTTGCTGCTGCGGTACAGCAGCAACAGGCTCCACCTTGCAGAAAGA 1043
QY 301 ThrHisArgAlaIleLeuPheCysIleGlnArgAspLeuProGlnAsnAsnAlaVal 320
DB 1044 ACACATGCTGCTATCTGTTTGGCCAGCAAGAAATTTGCTATCTCCAGCTAAGCAGTA 1103
QY 321 LeuValAlaSerGlyValAlaIleSerAsnPheThrIleArgArgAlaLeuGlnIleLeu 340
DB 1104 TTAGTGTGCTGAGAGGTGTGCAAGTAATCTTGTACATCCAGAGCATTTGAAATTTCTA 1163
QY 341 ThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeuCysThrAspAsnGly 360
DB 1164 GCAATGCAACACATGACCTTGTGTGCTCCCTCGAGACGTGACATGACATGAT 1223
QY 361 IleMetIleAlaIlePasnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAsp 380
DB 1224 ATCATGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1283
QY 381 IleGluGlyIleArgIleArgIleArgIleArgIleArgIleArgIleArgIleArg 399
DB 1284 GTAGAGACATCCATGAGAACCAAG 1310

RESULT 12

LOCUS AR428809 1526 bp DNA linear PAT 18-DEC-2003

DEFINITION Sequence 23 from patent US 6642041.

ACCESSION AR428809

VERSION AR428809.1 GI:40188595

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1526)

AUTHORS Chen,J., Feder,J.N., Nelson,T.C., Krystek,S.R. and Duclos,F.

TITLE Polynucleotides encoding a novel metalloprotease, MP-1

JOURNAL Patent: US 6642041-A 23 04-NOV-2003;

FEATURES

Location/Qualifiers

1. 1526

source

1. 1526

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.:	3.95e-108	Length:	1526
Score:	1385.00 <td>Matches:</td> <td>267</td>	Matches:	267
Percent Similarity:	100.00% <td>Conservative:</td> <td>0</td>	Conservative:	0
Best Local Similarity:	100.00% <td>Mismatches:</td> <td>0</td>	Mismatches:	0
Query Match:	65.18% <td>Indels:</td> <td>0</td>	Indels:	0
DB:	6 <td>Gaps:</td> <td>0</td>	Gaps:	0

US-10-649-273-2 (1-414) x AR428809 (1-1526)

QY 148 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnIysValGluPheProPheLeuVal 167

```

Db      1 ATGAGGCGTCATGACCTTACTATTAGGTTACCAATAAGTAGAATTTCTTTTAAAGTT
Oy      166 LeuLeuLeuSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu
Db      61 CTTTTCATTTCTTGAGAGCTCCTCTCTGTGGCATTAAGTTCAGAGAGTTTCAGATTTCCTG
Oy      188 LeuLeuGlyLysSerLeuAspPheLeuAlaProGlyAspMetLeuAspLysValAlaArgArg
Db      121 CTTCTTGGAAGAGCTTTGAGCATAGACCAAGTGACATGCTTGACAGAGTGGCAAGAGA
Oy      208 LeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyLysAlaIleGlnHis
Db      181 CTTTCTTTATATAAACAATCCAGAGTGTCTCCACCATGTGCTGGGAAGCCATAGAACAT
Oy      228 LeuAlaLysGlnGlyAsnAspPheHisPheAspPheLysProPheLeuHisHisAlaLys
Db      241 TTGGCCAAACAGAGAAATATATTTCAATTTGACATCAACCTCCCTTGACATCATCTTAA
Oy      248 AsnCyAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys
Db      301 AATTGTGATTTTCTTTTACTGATTCACACAGCTTATCTATATAATATATATATAAAG
Oy      268 GlyLysGlnGlnGlyIleGlnLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaAla
Db      361 GAAAAAGAGAGAGATTGAGAGAGGCAATCTGCTTCAGACAGACATTCGTGCC
Oy      288 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe
Db      421 ACAGTACACACACAATGAGATGTCATCTGTGTGAAGAACACATCGGCTATTTCTGTTT
Oy      308 CysLysGlnArgAspLeuProGlnAsnAlaValIleLeuValAlaSerGlyVal
Db      481 TGTAAACAGAGAGAGATTGTTACTCATAAATATATGCAATCTGTTGATCTGTGGTGTGTC
Oy      328 AlaSerAspPheThrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr
Db      541 GCAAGTACTCTCATATCCGACAGAGCTCTGGAATTTTAAACAACGACACAGCTGACCT
Oy      348 LeuLeuCyProProArgLeuCysThrAspAsnGlyIleMetIleAlaIleTPAsnGly
Db      601 TTGTTGTGTCCTCTCCACAGATATGCAATGATATGATGATATGATGATGATGATGATG
Oy      368 IleGlnArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGlnGlyIleArgTyrGln
Db      661 ATTGAAGAGCTACTGCTGCTGGGCAATTTTACATACATAGAAAGCATCTCCGCTATGAA
Oy      388 ProLysCyProLeuGlyValAspIleSerLysGlyValGlyGlnAlaSerIleLysVal
Db      721 CCMAATATCTCTTTGGAGTAGACATATCAAAAGAGTTGGAGAGCTTCATATAAAGTA
Oy      408 ProGlnLeuLysMetGluIle 414
Db      781 CCACAAATTAATAATGAGATTA 801

```

```

RESULT 13
BX934991      1522 bp      mRNA      linear      VRT 02-FEB-2004
LOCUS      Gallus gallus finished cDNA, clone CHEST189114.
DEFINITION      BX934991
ACCESSION      BX934991.1 GI:41635519
VERSION      BX934991.1
KEYWORDS      Gallus gallus (chicken)
SOURCE      Gallus gallus
ORGANISM      Gallus gallus (chicken)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 1522)
Boardman,P.B., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
Nibbelc,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
Tickle,C. and Wilson,S.A.

```

```

TITLE      Direct Submission
JOURNAL    Submitted (02-FEB-2004) Sanger Institute, Hinxton, Cambridgeshire,
           CB10 1SA, UK. E-mail enquiries: chickest@hms.umsf.ac.uk
COMMENT    BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
           sequencing project.
           This sequence is from the
           BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
           from a library constructed by Elizabeth Bosch. cDNA was prepared
           from RNA extracted from muscle, normalized, and poly A-tailed.
           EcoRI-NciI cut cDNA was then ligated into the vector. Vector:
           pBlueScript II KS(+); Site_1: EcoRI; Site_2: NciI Host: Escherichia
           coli DH10B.
FEATURES   Location/Qualifiers
           source          1..1522
                        /organism="Gallus gallus"
                        /mol_type="mRNA"
                        /strain="layer and broiler"
                        /db_xref="taxon:9031"
                        /clone="CHEST189114"
                        /clone_1db="CEORBN11"
                        /dev_stage="adult"
ORIGIN
Alignment Scores:
Pred. No.:      3,56e-106      Length:      1522
Score:          1362.00      Matches:      258
Percent Similarity:      83.20%      Conservative: 54
Best Local Similarity:    68.80%      Mismatches:   63
Query Match:      64.09%      Indels:      0
DB:              Gaps:      0
US-10-649-273-2 (1-414) x BX934991 (1-1522)
Oy      37 LysIleValLeuGlyIleGlnThrSerCysAspAspThrAlaAlaValAlaAspGlu
Db      254 AGACTTGTGCTGGGCAATCGAAACAGCTGTACACAGAGCGCGCTGCTGACAG
Oy      57 ThrGlyAsnValLeuGlyLysAlaIleHisSerGlnThrGlnValHisIleLysThrGly
Db      314 CGGGCACGCTGCTGGAGAGAGCCCTGCAGAGCCAAAGAGCTCACTCAAAAGAGT
Oy      77 GlyIleValProProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGln
Db      374 GGATTAATTCCTCATGATAGACACAGAGCTTCAAGAGAAAGATCCAGACAGTAAAG
Oy      97 GluAlaLeuSerAlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIle
Db      434 GAACCACTCAAGTGCAGAGGAGTTCTGTAATGAACTTGTGCTATTCGCAACTACAGTG
Oy      117 LysProGlyLeuAlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGly
Db      494 AATCCAGAGACTTGGCTGAGCTTGAAGTGGAGCTGACATCAAGCTTACAGCTGGTGAAC
Oy      137 GlnLeuLysLysProPheIleProIleHisHisMetGlnAlaHisAlaLeuThrIleArg
Db      554 AGTACCAAGAGAGCTTTTCATATCCCATTCATACATAGAGGCTCACAGCATACATCAGA
Oy      157 LeuThrAsnLysValGluPheProPheLeuValLeuLeuIleSerGlyGlyHisCysLeu
Db      614 CTGACAGAGCAAGATGATATTCCTCTTACTTCTTAACTCTCCGAGAGTACATGATC
Oy      177 LeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIleAla
Db      674 TTGGCAGTAGACACAGAGGTTTCAGATTTCTTGTGCTGACATTCATCATATATAGCA
Oy      197 ProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGluCys
Db      734 CCAGGTGATCATGCTGAGTATAGAGTACAGAGGCTCTTTTATGATAGACCCGAGTGC
Oy      217 SerThrMetSerGlyGlyValAlaIleGlnHisLeuAlaLysGlnGlyAsnArgPheHis
Db      794 CACGCAATGCTGGGGGAGAGGCAATAGACACCTGCTCAACCGAGAGCTGCGAACAG

```

Qy		237	PheAspiIleYsProPoleHniiShiSaIalysAenCySAsPheSePhethrGlyLeu	256			
Db		854	TACACTTTCAGACTCCCATGCACACAGATACGTACTGTGATTCTTTCTTCGGACTT	913			
Qy		257	GlnHiSValThrAspLysIleIlEmeLysLySGuLySGluGIUgLYlIecHuYsgLY	276			
Db		914	CAGAGCCTTGTCACAAAGCCATCTTCTTCAGAAAAGAAAAAGAAAGATTCAAGAAGG	973			
Qy		277	GlnIleLeuSerSerAlaAlaSpIleLaIaThrValGlnHiSthmeChIacyHis	296			
Db		974	GAATTCCTGTCCTCGTTTAAGACATGCTCTGCTGCACAGCACAGTGTGCTGCTCAT	1033			
Qy		297	LeuValYsaTgrTHriSaArgAlaileLeuPheCySLySGinArGaSPheLeuProGln	316			
Db		1034	ATTATCCAGCGGACACACCAGCCATGCTCTTCTTCAGAAAAACAGCATATTATTACA	1099			
Qy		317	AsnaSnlaValLeuValaIaSerGIyGIyValaIaSerAsmPheTyrlIeaRgaRga	336			
Db		1094	AAACTCGCACTCGTGTGTATCAGAGAGAGTGTGCAAGTAATCATATTACAGAAAAG	1153			
Qy		337	LeuGlulIleuThrAsnAlaThrgIncyshTrLeuLeuCySProProArgLeuCyS	356			
Db		1154	CTGCAAGCTCGGCAAAATGCAAAAGCGTTTGCTTTCTGTCTCTCCCTCCAGCGCTG	1213			
Qy		357	ThraSPasnGIyleMeIlEalaIatrpAsngLIyleGIuArgLeuArAlagLYleuGLY	376			
Db		1214	ACCATAATAGTGTATTATATTCATGCATGAATGCCATTTAAAGTTGCGTGCAGAGTGT	1277			
Qy		377	IleuhIsaSPIleGIuGIyIleArgTYrGIuProLyCySProlenGIyValaSPile	396			
Db		1274	ATTTATTCAGTACTAGTATGCATCGCATCCGATACGAACCAAAGCTCCCTTGGAATTAT	1333			
Qy		397	SerLySGluValIGlyGIuAlaSerIleLysValProGlnLeuLys	411			
Db		1334	TCCAAAAGAGTTGAAGAGATTTCCATCAAGTGCACCAAGACTAAC	1378			
RESULT 14							
BC038910							
LOCUS							
DEFINITION			Mus musculus O-sialylcoprotein endopeptidase-like 1, mRNA (CDNA				
ACCESSION			clone IMAGE:5053559), partial cds.				
VERSION			BC038910				
KEYWORDS			BC038910.1 GI:24433548				
SOURCE							
ORGANISM			Mus musculus (house mouse)				
REFERENCE							
AUTHORS			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1017) Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Butow, K.H., Schaefer, C.F., Bat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hasteh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stepleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schneerz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carinciul, F., Prange, C., Rada, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., Mcwan, P.J., McNernan, K.C., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Philky, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kerteman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Merris, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 22388257 JOURNAL MEDLINE PUBMED 12477932 REFERENCE 2 (bases 1 to 1017)				

FEATURES	source
ALIGNMENT SCORES:	<p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: http://image.lnl.gov</p> <p>Series: IRAK Plate: 86 Row: F Column: 12.</p> <p>Location/Qualifiers</p> <p>1. 1017</p> <p>/organism="Mus musculus"</p> <p>/mol_type="mRNA"</p> <p>/strain="FVB/N"</p> <p>/db_xref="taxon:10090"</p> <p>/clone="IMAGE:5053559"</p> <p>/tissue_type="Liver, normal, 5 month old male mouse."</p> <p>/clone_lib="NCI CGAP_Lig"</p> <p>/lab_host="DH10B"</p> <p>/note="Vector: pCMV-SPORT6"</p>
ORIGIN	
Alignment Scores:	
Pred. No.:	7.91e-99
Score:	1273.00
Percent Similarity:	92.73%
Best Local Similarity:	87.64%
Query Match:	59.91%
DB:	10
US-10-649-273-2 (1-414) x BC038910 (1-1017)	
QY	140 LysPProPheIleProIleHisMetGluAlaHisAlaLeuThrIleArgLeuThrAsn 159
DB	2 AAGCATTCATCCGATTCATCATCATGAGGCTCAGCAGCTGACTATTAAGGCTCACCAAT 61
QY	160 LysValGluPheProPheLeuValLeuLeuIleSerGlyGlyHisCysLeuAlaLeu 179
DB	62 AAAGTAGAATTTCTTTTATGTTCTTTGATTTCTGGGGCTACGCTGTTGGGATTA 121
QY	180 ValGlnGlyValSerAspPheLeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAsp 199
DB	122 GTCCAGAGGTGTTCCGATTTCTGCTCTTGGGAAGCTCTTGACATACACAGGCGAC 181
QY	200 MetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGluCysSerThrMet 219
DB	182 ATGCTGTGCAAGGTGGCAAGAAGATTTCTTTAATCAACATCCAGAAATGTTCTACAATG 241
QY	220 SerGlyGlyLysAlaIleGluHisLeuAlaLysGlnGlyAsnArgPheHisPheAspIle 239
DB	242 AGTGGTGMAAAGCATATGAAACAGTGGCCAAAGCGMAAATGATTCCATTTTACTATC 301
QY	240 LysProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGlyLeuGlnHisVal 259
DB	302 AATCCACTATGACAGAAATGCTAAGAAATGCGATTTTCTTTACCGGACCTTCAACATAT 361

Oy		260	ThrsApIleatIIeMeLyIsGslusgluglulgylilegllwlysgLGIHlleu	279
Dd		362	ACTATAAGCTAAATPACACACAAGAANNANAGAGCATTTGGAGAGGCCAATTCTTG	421
Oy		280	SerSerAlaaIasppIlleaIalaIthaValGlInhlstRmetAlaCyshlsleuVallys	299
Dd		422	TCA TCAGGTGCAGACATTCGTC GCGCSTA CAGCATGCACA AGCGTGCAC TTAGCCGAAA	481
Oy		300	ArgThrHisahgaLaalleuePhcGalygslnlrgapbleuleuProglinsAsnMa	319
Dd		482	AGAACACATCCCGCTAATCTCTTTTTGCAMAGCAAAMAATTTGCTCTCCAGCTTAACGA	541
Oy		320	VallenuAlalasegLIgLYVAlalaseraNpheTyrlleaYrGaPaIaleuGlulle	339
Dd		542	GTA TTAGTGTAATCTGGAGGtggTGCAAGTA ACTTGTACATCCGAAAAAGCATTTGAAAT	601
Oy		340	LethTrasnAlathngIncysTheLeudeuCysPropoProArgleucycsThraspaSn	359
Dd		602	GTCCC AATGACACGCACATGCAC GTTGTGTGCACC CTCCAAGACTGTGCAC TGACAT	661
Oy		360	GlylleMetllaAtTPASNgLYlleGLuargleuaIaglaAglyleuGlylleuHls	379
Dd		662	GGCATCATGATTCGATGCATGAATGSAATTGAAGAATTACCGTCGTCGCGGCTTTACAT	721
Oy		380	AspllegluIgLYlleargTYrghlpolyGscypsoleuGlyVAlaspIIlesetIrysgS	399
Dd		722	GATGTAGAAGACATCCGATGATGAACCAA AAAATGCTCTTGAGATGNACATATCCAGAGA	781
Oy		400	ValglyGUAlaserilleyelvalproglneulusmectGuille 414 ::: ::: :	
Dd		782	GTTCGAGAGCTGCATTAAGTACCGCATTTAAAAAATGGCAC ATT 826 ::: ::: :	
RESULT 15	BX930963	1558 bp	mRNA linear	VRT 30-MAR-2004
LOCUS	BX930963			
DEFINITION	Gallus gallus finished cDNA, clone CHEST62nl6.			
ACCESSION	BX930963			
VERSION	BX930963.2 GI:46016890			
KEYWORDS				
SOURCE				
ORGANISM	Gallus gallus (chicken)			
AUTHORS	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archaeopteria; Aves; Neognathae; Galliformes; Phasianidae;			
REFERENCE JOURNAL	Phasianinae; Gallus. Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.B., Crooking,M.D.R., Davies,R.M., Francis,M.D., Graham,D.V., Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison M., McLaren,S.R., Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G., Tickle,C. and Wilson,S.A.			
COMMENT	Direct Submission Submitted (29-MAR-2004) Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: chickest@bm.unist.ac.uk On Apr 1, 2004 this sequence version replaced gi:41631491. BBSCR/Dundee/Nottingham/Sanger/Sheffield/UWIST Gallus gallus cDNA sequencing project. This sequence is from the BBSRC/Dundee/Nottingham/Sanger/Sheffield/UWIST cDNA collection, from a library constructed by Elizabeth Bosch. cDNA was prepared from RNA extracted from limbs, and poly A-tailed. EcoRI-NotI cut cDNA was then ligated into the vector. Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia coli DHIOB.			

FEATURES
source

```
/organism="Gallus gallus"  
/mol_type="mrna"  
/strain="White leghorn, Hisex"  
/db_xref="taxon:9031"  
/clone_name="CHEST62n16"  
/clone_1b="CSEBCHL20"  
/dev_stage="stage 36"
```

ORIGIN

Alignment Scores:	
Pred. No.:	4-57e-93
Score:	1208.00
Percent Similarity:	82.54%
Best Local Similarity:	67.46%
Query Match:	56.85%
Gap:	0
Indel:	0
Mismatch:	0
Match:	1558
Length:	1208

US-10-649-273-2 (1-414) X BX930963 (1-1558)

QY 74 LysThrGlyGlyIleValProProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArg 93

Db 486 AGAGCAGGTGGAATAATTCCTCATGTAGCACAGCAGCTTTCACAGAGAAAGCAATCCAGCAA 545

```

94 l l e v a l g i n g l u a l e u s e r a l a s e r g l y w a l s e r p r o s e r a b p l e u s e r m i d l e a d i l s
  . . . . . : : : : : .

```

[illegible][illegible]

100

[illegible]

154 *Abstracts of the 1973*

726 A C C A T T C A G A C C A G C C A G T A G A T T T T C C C T T C T T A C T C T C C G A G C T 785

174 HisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuLeuGlyLysSerLeu 193

Db 786 CACTGCATCTTGGCAGTAGCACCAGGAGTTTCAGATTTCCTTCTGCTTGGACAGTCCATA 845

QY 194 AspIleAlaProGlyAspMetLeuAspIlyValAlaArgArgLeuSerLeuIleLysHis 213

Db 846 GATATAGCACCAGGTGACATGCTGGATAAGGTAGCAAGAAGCTCTCTTTAGTGAAGCAC 905

214 ProglucySerThrMetSerGlyGlyAlaIleGluHisLeuAlaGlyGlnGlyAsn 233

Db 906 CCGAGTGCACCGCA1GGCTGGGGGAAGCCAA1AGAGCACCTGGCTCAACCGAGAC 965

234 AIGPHENIBPNEAVRIELVBPFOIODEWHBIBAIADYBASINCYBABPFNECEIFNE

[illegible][illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

2004 年 11 月 26 日

1145 GCTGCTTATATATCCAGGACACACCGAGCCATGCTCTTCTGCATGAAACAGCATA 1200

314 LeuProGlnAsnAsnAlaValLeuValAlaSerGlvGlValAlaSerAsnPheTyrIle 333

Db 1206 TTATTACCAAACTGCACTCTGGTGTATCAGGAGGAGTTCGAAGTAATCAGTATATC 1266

334 ArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProPro 353

Db 1266 AGAAAGACTGCAGACTCTGGCAATGCAACGGTTTTCGCTTTCTGTCTCCTCTCCA 132

354 ArgleuCystrAspaNGLyIleMetIleAlaTPAsNGLyIleGluArgleuArgAla 373

Db 1326 AGCGTGTCCACCGATAATGGTGTATGATTCATGGAAATGGCAATGAAAGGTTCCTGCA 1336

[illegible]

```

Db      1386  GGATGCTGTAATTTATACAGTACTGATGGCATCGCTATGACCAAGCTCCCTTGA 1445
Oy      394  ValAspIleSerIysGluValGlyGluAlaSerIleIysValProGlnLeuLys 411
Db      1446  ATTGATATTTCCAAAAGAGTTGAAAGGATTCATCCAGTGGCCAAAGACTTAAG 1499

```

Search completed: June 16, 2005, 23:18:30
 Job time : 4983.35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 16, 2005, 12:03:59 ; Search time 598.5 Seconds
(without alignments)
4094.859 Million cell updates/sec

Title: US-10-649-273-2
Perfect score: 2125
Sequence: 1 MLTKTAGVFPKPSKRYK...DISKEVGEASIKVPLMEI 414

Scoring table: BLOSUM62
Xgapop 10.0, Ygapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Ygapext 7.0
Delop 6.0, Delext 7.0

Searched: 4390206 seqs, 295870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+.p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10649273/runat.15062005.111415.6022/app.query.fasta_1.1429
-DB=N Geneseq.16Dec04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10649273.cgn2.1.1.1063 @runat.15062005.111415.6022 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq.16Dec04:.*
1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002as:.*
7: geneseqn2002bs:.*
8: geneseqn2003as:.*
9: geneseqn2003bs:.*
10: geneseqn2003cs:.*
11: geneseqn2003ds:.*
12: geneseqn2004as:.*
13: geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2125	100.0	2058	6 ABA93268 Human O-s
2	2125	100.0	2197	6 ABS76635 DNA encod
3	2125	100.0	2572	8 ABT23207 Human pro
4	2088	98.3	1820	6 AAd46856 Human g1y
5	2088	98.3	1820	8 ACA60887 Human cDN

6	2088	98.3	1821	10 ABS57020	ABs57020 cDNA enco
7	1944	91.5	2208	10 ADAS2832	Ada52832 Human cod
8	1944	91.5	2890	12 ADQ24627	Adq24627 Human sof
9	1747	82.2	1416	6 ABX70950	Abx70950 Novel hum
10	1549	72.9	1572	6 ABQ75508	Abq75508 Murine si
11	1385	65.2	1526	6 ABS76639	AbS76639 DNA encod
12	1001.5	47.1	3358	10 ADB31345	ABs31345 Human dia
13	714.5	33.6	1601	4 ABL24633	ABl24633 Drosophil
14	700.5	33.0	1385	4 AAH15110	AAh15110 Human cDN
15	687.5	32.4	3656	4 ABL24632	ABl24632 Drosophil
16	662.5	31.2	1557	3 AAC38454	AAc38454 Arabidops
17	629	29.6	571	4 AAH12983	AAh12983 Human cDN
18	599	28.2	2734	5 AAS84622	AAs84622 DNA encod
19	531.5	25.0	4360	6 AAD48239	AAd48239 Ehrlichia
20	524	24.7	1146	8 ACA26804	ACa26804 Prokaryot
21	502	23.6	1044	8 ACA39102	ACa39102 Prokaryot
22	502	23.6	94750	4 AAF28551	AAf28551 Genomic f
23	501	23.6	1053	12 ADL03120	ADl03120 DNA encod
24	500	23.5	1092	6 ABQ90383	ABq90383 M. capuli
25	499.5	23.5	1092	6 ABD02197	ABd02197 Pseudomon
26	494.5	23.3	1032	8 ACA43173	ACa43173 Prokaryot
27	492.5	23.2	1026	4 AAS54064	AAs54064 Pseudomon
28	492.5	23.2	1026	8 ACA42146	ACa42146 Prokaryot
29	492.5	23.2	1026	10 ADG73341	ADg73341 P aerugin
30	492.5	23.2	1059	11 ABD02280	ABd02280 Pseudomon
31	490.5	23.1	1026	10 ADG73343	ADg73343 P aerugin
32	484.5	22.8	1014	4 AAS52570	AAs52570 E. coli D
33	484.5	22.8	1014	8 ACA51431	ACa51431 Prokaryot
34	484.5	22.8	1014	8 ACA32641	ACa32641 Prokaryot
35	483.5	22.8	1014	8 ACA54107	ACa54107 Prokaryot
36	482.5	22.7	1020	8 ACA44384	ACa44384 Prokaryot
37	482.5	22.7	1074	10 ADF02056	ADf02056 Bacterial
38	481.5	22.7	1014	4 AAS56045	AAs56045 Salmonell
39	479	22.5	1315	2 AAQ27645	AAq27645 Glycoprot
40	475.5	22.4	1029	4 AAS53309	AAs53309 Haemophil
41	475.5	22.4	1029	8 ACA34150	ACa34150 Prokaryot
42	475.5	22.4	110000	2 AAT42063	05 Continuation (6 of
43	474.5	22.3	1032	11 ACH96255	ACh96255 Klebsiell
44	472.5	22.2	1014	10 ACF71364	ACf71364 Phototrab
45	472.5	22.2	110000	10 ACF7367_45	Continuation (46 o

ALIGNMENTS

RESULT 1	
ABA93268	
ID	ABA93268 standard; cDNA; 2058 BP.
XX	
AC	ABA93268;
XX	
DT	19-APR-2002 (first entry)
XX	
DE	Human O-sialoglycoproteinase-like protein encoding cDNA SEQ ID NO:1.
XX	
KM	Human; O-sialoglycoproteinase-like protein; OSGLP; enzyme; gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	/*tag= a
FT	/product= "O-sialoglycoproteinase-like protein"
XX	
PN	CN118550-A.
XX	
PD	24-OCT-2001.
XX	
PF	19-APR-2000; 2000CN-00106834.
XX	
PR	19-APR-2000; 2000CN-00106834.
XX	
PA	(SHAN-) SHANGHAI SHENGYUAN GENB DEV CO LTD.
XX	

PI Mao Y, Xie Y;
 DR WPI: 2002-115090/16.
 DR P-PsDB; ABB05481.
 XX
 XX O-sialoglycoproteinase-like protein and encoding polynucleotide, useful
 PT for diagnosing, preventing and treating related diseases.
 PR
 XX
 XX Claim 5; Page 29-30 (Disclosure); 38pp; Chinese.
 PS
 XX
 XX The present sequence encodes human O-sialoglycoproteinase-like protein
 CC (OSGPRP). The present invention also describes: (1) the preparation of
 CC the OSGPRP protein; (2) applying the OSGPRP protein in diagnosis; (3) the
 CC prevention and/or treatment of related diseases; (4) utilizing the OSGPRP
 CC protein in screening its agonist, excitomotor and inhibitor and preparing
 CC an antibody against the OSGPRP protein; and (5) the use of the OSGPRP
 CC polynucleotide sequences, proteins, agonists, excitomotors, inhibitors
 CC and antibodies in treating diseases related to the abnormal OSGPRP gene
 CC and in preparing the medicine composite for the treatment
 XX
 SQ Sequence 2058 BP; 637 A; 400 C; 410 G; 611 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,94e-212 Length: 2058
 Score: 2125.00 Matches: 414
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-649-273-2 (1-414) x ABA93268 (1-2058)
 QY 1 MetLeuIleLeuThrIysThrAlaGlyValPhePheLysProSerLysArgLysValTyr 20
 Db 110 ATGTAATTCCTTCGACATACAGCTGACGAGAGTTTAAAAACATCAAAAAGAAAGTTAT 169
 QY 21 GluPheLeuArgSerPheAsnPheHisProGlyThrLeuPheLeuHisLysIleValLeu 40
 Db 170 GAATTTTAAAGAGTTTAAATTTCACTTCCTGGAACACTATTTCTTCAATAAAATAGAAATTG 229
 QY 41 GlyIleGluThrSerCysAspAspThrAlaAlaValValAspGluThrGlyAsnVal 60
 Db 230 GGAATTTGAAACTAGTTGTGATGATACAGCAGCTGCTGTGTGATGATGAATCTGAAATGTG 289
 QY 61 LeuGlyGluAlaIleHisSerGlnThrGluValHisLeuLysThrGlyIleValPro 80
 Db 290 TTGGGAGAAACAATACATATCCCAACTGAAGTTCAATTAATAAAACAGGTGGATTGTCT 349
 QY 81 ProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGluAlaLeuSer 100
 Db 350 CCAGACGCTCAACAGCTTCACAGAGAAATATTCACGAAATATACACAGAGCTCTTTCT 409
 QY 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleLysProGlyLeu 120
 Db 410 GCCAGTGAAGTCTCTTCAAGTGAACCTCTCAGCAATGCAATACCAATAAACAAGACCT 469
 QY 121 AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuLysLys 140
 Db 470 GCTTTAAGCTGGAGGTGGCTGCTTAATTAAGCTTACAGCTGTATGACACGTTAAAAAAG 529
 QY 141 ProPheIleProIleHisIleMetGluAlaHisAlaLeuThrIleArgLeuThrAsnLys 160
 Db 530 CCATTCATTCATCCATTCATCATATGAGAGCTCATGCACTTACTATTAGGTGACCAATAAA 589
 QY 161 ValGluPheProPheLeuValLeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuVal 180
 Db 590 GTAGAATTCCTCTTTTAACTTCTTTTGATTCTGGAGGTCAAGTCTGTTGGCAATTAAGT 649
 QY 181 GlnGlyValSerAspPheLeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMet 200
 Db 650 CAGAGAGTTTCAGATTTTCTGCTCTTGGAAAGACTTTTGGACATAGCACAGGTGACATG 709
 QY 201 LeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGluLysSerThrMetSer 220

Db	710	CTTGACAAAGTGGCAAGAAAGACTTCTTTAATATAAAGATCCAGAGTGCTCCACCATGAGT	769
Oy	221	GLYGLYLVALAILEGLUHSLEUALALYSGINGLYASNAARGPHEHSI	240
Db	770	GGTGGGAAGCCATGAAACATTGGCCCAAAACAAGAAATGATTTTCATTATTTGACATCAA	829
Oy	241	PROPRLEUHSIHIALALYSASNCYASAPHPESERPHETHGLYLEUGINHISVAL	260
Db	830	CCTCCCTTGACATCAGCTPAAAAATTTGATTTTCTTTACTGAGACTTCACACGTTACT	889
Oy	261	ASPLYSILEILEMETELYSGLULYSGIUGIUGIYLEGLULYSGLYGINILEUSER	280
Db	890	GATATAAATPATATATGAAAAAGCAAAAAGAGAGATTTGAGAGGGGCAAACTCTGTCT	949
Oy	281	SERIALAALASPILEALALATHRVALGINHSITRMETALCYSHISLEUVALYSARG	300
Db	950	TCAGACAGACATGCTGCTCCACAGTACGACACACATGGCAATGTCATCTTGTGAAAAGA	1000
Oy	301	THRHSARGALAILLEUPHESYSGINARAPHEULEUPROCLINENASENALAVAI	320
Db	1010	ACACATCGGGCTATTCTGTTTGTGTAAGACAGAGACTGTGTACTTCAAAATTAATGACATA	1060
Oy	321	LEUVALIASERGLYGLYVALALASERASNPHERYILLEARGARGALAEUGIUILEU	340
Db	1070	CTGGTTGCATCTGGAGGTGTCCCAAGTAACTTCTATATCCGACAGCTCTGGAAATTTTA	1120
Oy	341	THRASNALATHRGLNCYESTHRLEULEUCYSPROPROAARGLEUCYTHRAASANGLY	360
Db	1130	ACAAACGACACACAGCTGACCTTGTGTGTCTCTCCACAGACTATGACACTGATAATATGAC	1180
Oy	361	ILEMETTLEALATPASNGLYILEGLUARGLEUARGALIGLYLEUGIYILLEUHSIASP	380
Db	1190	ATTATGATATGTCATGGAATGTATGTAAAGACTACACGTGCTGGCGATTTTATCAATGAC	1240
Oy	381	ILEGLUGIYILEARYTRGULUPROLYSCYSPROLEUGLYVALASPILESERLYSGIUAL	400
Db	1250	ATGAGAGGATCCGCTATGATGACCAAAATGTCTCTTGAGCTAGACATATCAAAAGANGTT	1300
Oy	401	GLYGLUALASERILELYSVALPROGLINLEULYSMETGLUILE	414
Db	1310	GSAGAGGCTTCATTAAGTATGACACATTTAATAAATGAGATATA	1351
RESULT 2			
ABST76635			
ID	ABST76635	standard; DNA; 2197 BP.	
XX	AC	ABST76635;	
XX	DT	11-DEC-2002 (first entry)	
XX	DE	DNA encoding novel human metalloprotease MPI.	
XX	KM	Metalloprotease; MP-1; immune disorder; glutamate transport; cancer;	
XX	KM	motor neuron disorder; amyotrophic lateral sclerosis; ALS; diabetes;	
XX	KM	reproductive disorder; Kleinfelter's syndrome; germinal cell aplasia;	
XX	KM	genital wart; metabolic disorder; premature puberty; Kallman syndrome;	
XX	KM	Cushing's syndrome; neurodegenerative disease; Alzheimer's disease;	
XX	KM	Parkinson's disease; Huntington's disease; Tourette syndrome; sepsis;	
XX	KM	liver disease; renal disease; immune disorder; rheumatoid arthritis;	
XX	KM	acquired immunodeficiency syndrome; AIDS; pulmonary disease; pneumonia;	
XX	KM	emphysema; cystic fibrosis; vascular disorder; inflammatory disorder;	
XX	OS	neurological disorder; gene; ds.	
XX	OS	Homo sapiens.	
XX	PN	WO200272751-A2.	
XX	PD	19-SEP-2002.	
XX	PP	05-FEB-2002; 2002MO-US003353.	
XX	PR	05-FEB-2001; 2001US-0266518P.	

PR 10-APR-2001; 2001US-0282814P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Chen J, Feder J, Nelson TC, Duclos F, Krystek S,
XX
WPI: 2002-723329/78.
DR P-PSDB; ABG96478.
DR

PT New isolated nucleic acid encoding MP-1 protein, useful for preventing
PT treating, or ameliorating diseases associated with aberrant
PT metalloproteinase activity, e.g. immune, inflammatory and
PT neurological disorders.

Claim 1; Fig 1A-C; 473pp; English.

The invention describes an isolated nucleic acid molecule (I) encoding metalloproteinase (MPL-1). (I) is useful for preventing, treating, or ameliorating a medical condition, particularly an immune disorder, an aberrant glutamate transport or motor neuron disorder, such as amyotrophic lateral sclerosis (ALS), its juvenile form or an ALS-like condition. The compositions and methods are also useful for diagnosing, prognosticating, treating, ameliorating and/or treating disorders associated with Mpl-1 activity, e.g. diabetes, cancer, reproductive disorders (e.g. Klinefelter's syndrome, genital warts, or germinal cell aplasia), metabolic disorders (e.g. premature puberty, Kallman syndrome, or Cushing's syndrome), neurodegenerative diseases (Alzheimer's disease, Parkinson's disease, Huntington's disease or Tourette syndrome), liver and renal diseases and immune disorders (e.g. AIDS, rheumatoid arthritis or sepsis), pulmonary diseases (e.g. pneumonia, emphysema or cystic fibrosis) and vascular, inflammatory and neurological disorders (e.g. Alzheimer's disease or Parkinson's disease). This sequence represents a metalloproteinase Mpl polynucleotide

Sequence 2197 BP; 681 A; 441 C; 439 G; 636 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.14e-212	length:	2197
Score:	2125.00	Matches:	414
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0

US-10-649-273-2 (1-414) X ABS76635 (1-2197)

QY	1	MetLeuIleLeuThrIleThraIaGlyValAlphePheLysProSerLysAaGlyValAltyr	20
Db	231	ARGCTNATCTGACTGACGAGAGTTTTTTTTTAAACATCAAAAAGAAAGTTTAT	290
QY	21	GIuPheLeuAaRSeRpheAaSPheNiSPROGlyThrLeuPheLeuNiSPySIIeValLeu	40
Db	291	GAATTTTTTAAGAAGTTTTTAAATTTTCATCCGGAACCATATTTCTCATAAATAAGATTTG	350
QY	41	GlyIIeGIuThSeRCysAspSphThraIaAlaIaValAlaSPGIuThrgIyaSPVal	60
Db	351	GGAATTTGAAACTAGATTGTGTATGATATACAGCAGCTGCTGTGTGATCAAACTGAAATATG	410
QY	61	LeuGIyGIuIaIIeHisSerGIuThrgIuValHiSPeLyuThrgIyGIyIIeValPro	80
Db	411	TTGGGGAAGCAATATACATTTCCAAACTGAAAGTCATTTAAAAAACAGGTGGAAATGTTCT	470
QY	81	ProAlaAlaGIuNIleuHiSPArgGIuAsnIIeGIuAaRgIIeValGIuGIuIaIleuSer	100
Db	471	CCAGCAGCTCAACAGCTTCACAGAGAAANATTCAAGAAATATATACAGAGAGCTTTCT	530
QY	101	AlaSPeGIyValISeRProSeRaSPeLeuSeRaIIeAlaThrThriIeLysPROGIyLeu	120
Db	531	GCACGTGGAGTCTCTCCAAGTGACCTCTCAGCAATTGCCAATCCATATAAACCGACGACTT	590
QY	121	AlaIleuSerLeuGIyValGIyLeuSerPheSerLeuGIuIleValGIyGIuIleLysIle	140
Db	591	GCITTTAAGCTCGAGATGGGCTTATATCATTTAGCTTACAGCTGTGTAGACAGTTTAAAAAG	650

Accession	Gene	Protein	Position	Sequence	Length
QY	Probel	Probel	161	Probel	160
Db	651	CCATTTCATCCCATTCATCATATGAGGCTATCATCTTACTATTAGGTTGACCAATPAA	710		
QY	161	ValGluPheProPheLeuValLeuLeuIleSerGlyGlnHisCysLeuLeuAlaLeuVal	180		
Db	711	GTAAGAATTCCTTTTATTTAGTTCTTTTGATTTCTGAGAGTCACTGTCTGTGGCATTTAGTT	770		
QY	181	GlnGlyValSerAspPheLeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMet	200		
Db	771	CAAGAGATTCAGATTTTCTGCTCTTGGAAGAATCTTTGACATATGACACGAGTGACATG	830		
QY	201	LeuAspLysValAlaIaIaGArgLeuSerLeuIleLysHisProGluCysSerThrMetSer	220		
Db	831	CTTGACAAGAGTGGGACAGAGAATCTTTTAAATAAACAATCCAGAGTGTCCACATAGAT	890		
QY	221	GlyGlyLysValAlaIleGlnHisLeuAlaLysGlnGlyAsnArgPheHisAspAspIleLys	240		
Db	891	GGTGGGAAGGCATATGAAACATTTGGCCCAACAGAGAAATTAATTTCAATTTTGACATCAA	950		
QY	241	ProPheLeuHisSerAlaLysAsnCysAspPheSerPheThrGlyLeuGlnHisValThr	260		
Db	951	CCTCCCTTGATCATGCTCTTAATAATTTGATTTCTTTTACTGACATTCACACAGTTACT	1010		
QY	261	AspLysIleIleMetLysLysGlnLysGlnGlyIleGluLysGlyGlnIleLeuSer	280		
Db	1011	GATTAATAATATATGAAATGAAAGAAAGAGAGATGATTGAGAGGGGCAATCTCTGCT	1070		
QY	281	SerAlaAlaAspIleAlaIleThrValGlnHisThrMetAlaCysHisLeuValLysArg	300		
Db	1071	TCAGCAGACAGACATTTGCTGCCACAGTACACAGACCAAGGAGATGTCATCTTGTAAGAA	1130		
QY	301	ThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaVal	320		
Db	1131	ACACATCGGGCTATTCGTGTTTGTATGACAGAGACCTTGTTACTTCANAAATATATGCACTA	1190		
QY	321	LeuValAlaSerGlyGlyValAlaSerAsnPheTyrIleArgArgAlaLeuGluIleLeu	340		
Db	1191	CTGTGTGATCTGTGTGTGTGTGCGCAAGTACTTCTATATCCGACAGAGCTCTGGAAATTTTA	1250		
QY	341	ThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeuCysThrAspAsnGly	360		
Db	1251	ACAAACGACACACAGTGACCTTTGTGTGTCTCTCCCAACATATGACATGATTAATGAC	1310		
QY	361	IleMetIleAlaTyrAsnGlyIleGlnArgLeuArgAlaGlyLeuGlyIleLeuHisAsp	380		
Db	1311	ATTATGATTTGATGAAATGATGATTTGAAAGATCTAGCTGTGGCATTTTACATGAC	1370		
QY	381	IleGluGlyIleArgTyrGluProLysCysProLeuGlyValAspIleSerLysGluVal	400		
Db	1371	ATGAGAGGCATCCGCTATGAAACCAAAATGCTCTTGAGATGACATATCAAAAGAAAGTT	1430		
QY	401	GlyGluAlaSerIleLysValProGlnLeuLysMetGluIle	414		
Db	1431	GGAGAGCTTCATTAAGTACACCAATTAATAAATGAGATTA	1472		
RESULT 3					
ABT23207					
ID	ABT23207	standard; DNA; 2572 BP.			
AC	ABT23207;				
XX					
DT	01-MAY-2003	(first entry)			
XX					
DE		Human protein modification + maintenance molecule DNA SEQ ID No 36.			
XX					
KW		Cytostatic; antiarteriosclerotic; anti-HIV; antiallergic; anticonvulsant;			
KW		cardioprotective; antiparkinsonian; nootropic; antituberculous;			
KW		antitumor; hepatotropic; gynecological; antibacterial; vitamin;			
KW		procoagulant; antiparasitic; cell proliferative disease; PMOD;			
KW		protein modification and maintenance molecule; immunogenic fragment;			
KW		cancer; autoimmune; inflammatory disease; neurological disorder;			

	RESULT 3
ABR23207	
ID	ABT23207 standard; DNA; 2572 BP.
XX	
AC	
XX	ABT23207;
XX	
DT	01-MAY-2003 (first entry)
XX	
DE	
XX	Human protein modification + maintenance molecule DNA SEQ ID No 36.
XX	
KW	Cytostatic; antiarteriosclerotic; anti-HIV; antiallergic; anticonvulsant;
XX	cerebroprotective; antiparkinsonian; nootropic; antiinflammatory;
KW	antitumor; hepatotropic; gynaecological; antibacterial; virucide;
KW	prozoacide; antiparasitic; cell proliferative disease; PMOD;
	protein modification and maintenance molecule; immunogenic fragment;
KW	cancer; autoimmune; inflammatory disease; neurological disorder;

KW gastrointestinal; developmental; vesicle trafficking disorder; infection;
 KW protein-protein interaction; drug-target interaction;
 KW gene expression profile; human; gene; ds.

OS Homo sapiens.

XX WO200300844-A2.

XX 03-JAN-2003.

PF 18-JUN-2002; 2002WO-US019360.

XX 22-JUN-2001; 2001US-0300508P.

PR 06-JUL-2001; 2001US-0303445P.

PR 13-JUL-2001; 2001US-0305405P.

PR 09-AUG-2001; 2001US-0311442P.

PR 24-AUG-2001; 2001US-0314821P.

PR 29-AUG-2001; 2001US-0315922P.

PR 03-MAY-2002; 2002US-0378205P.

XX (INCY-) INCYTE GENOMICS INC.

XX Gandhi AR, Kable AE, Swarnakar A, Hafalia AJA, Tran B, Duggan BM;

PI Warren BA, Ison CH, Honchell CD, Nguyen DB, Lu DAM, Lee EA, Yue H;

PI Forsythe JD, Barroso I, Ramkumar J, Griffin JA, Li JX, Yang J;

PI Thangavelu K, Gietzen KJ, Ding L, Baughn MR, Borowsky ML, Yao MG;

PI Wala NK, Mason PM, Gururajan R, Lee S, Becha SD, Lee SY, Tran UK;

PI Elliott VS, Luo W, Sprague KW, Tang YT, Lu Y, Zebardadian Y;

XX WPI; 2003-184039/18.

DR P-PSDB; ABJ26654.

XX New isolated human PMOD polypeptide and polynucleotide, useful for

PT diagnosing, treating and preventing diseases or conditions associated

PT with the aberrant PMOD expression, e.g. cancer, AIDS, atherosclerosis and

PT infections.

XX Claim 91; Page 211; 225pp; English.

XX The invention relates to an isolated polypeptide comprising: any of 28
 CC sequences of 48-1256 amino acids; a natural amino acid sequence at least
 CC 90% identical to the 28 amino acid sequences, 94% identical to a sequence
 CC of 703 or 267 amino acids, 96% identical to a sequence of 414 amino
 CC acids, or 97% identical to a sequence of 242 amino acids, all given in
 CC the specification, or a biologically active or immunogenic fragment of
 CC the isolated polypeptide. The polypeptides and polynucleotides are useful
 CC in diagnosing, treating and preventing diseases or conditions associated
 CC with the decreased expression of protein modification and maintenance
 CC molecules (PMOD), such as cell proliferative diseases (e.g. cancer,
 CC atherosclerosis), autoimmune/inflammatory diseases (e.g. AIDS,
 CC allergies), neurological disorders (e.g. stroke, Parkinson's disease,
 CC epilepsy), gastrointestinal (e.g. ulcer, cirrhosis), reproductive (e.g.
 CC endometriosis), developmental, vesicle trafficking disorders, and
 CC infections (e.g. bacterial, viral, parasitic, protozoal). These are also
 CC useful in assessing the effects of exogenous compounds on the expression
 CC of nucleic acid and amino acid sequences of PMOD. The PMOD or its
 CC fragments are useful in screening compounds for effectiveness as agonist
 CC or antagonist of the polypeptides, or in altering the expression of the
 CC target polynucleotide and compounds that specifically bind to or modulate
 CC the activity of the polypeptide. The microarray is useful in monitoring
 CC or measuring protein-protein interactions, drug-target interactions, and
 CC gene expression profiles. This polynucleotide sequence represents the DNA
 CC encoding a human PMOD protein of the invention

XX Sequence 2572 BP; 780 A; 489 C; 509 G; 794 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 2.7e-212 Length: 2572

XX Score: 2125.00 Matches: 414

XX Best Local Similarity: 100.00% Conservative: 0

XX Query Match: 100.00% Mismatches: 0

XX DB: 8 Gaps: 0

US-10-649-273-2 (1-414) x ABT23207 (1-2572)

QY	1	MetLeuIleuThrIysThrAlaGlyValPhePheIysProSerIysArgIysValTyr	20
DB	144	ATGCTAATCTTGACTTACGACGAGAGTTTATTTTAAACATCAAAAGAAAGTTTAT	203
QY	21	GIuPheLeuArgSerPheAsnPheHisProGlyThrLeuPheLeuHisIysGileu	40
DB	204	GAATTTTAAAGAGTTTAAATTTTCACTCGGAACACATATTTCTTCAATAAATGATATG	263
QY	41	GlyIleGluThrSerCysAspAspThrAlaAlaValValAspGluThrGlyAsnVal	60
DB	264	GGAATTAACCTAGTTGTGATGATACAGACAGCTCTGTGTGATGAACCTGGAATATG	323
QY	61	IeuGlyIuAlaIleHisAsenGlnThrGluValHisIleuIysThrGlyIleValPro	80
DB	324	TTGGAGACGACAAATACATTCCTCAAACTGAAGTTTATTTAAACAGGAGGATTTCT	383
QY	81	ProAlaIaGlnGlnIleuHisArgGluAsnIleGlnArgIleValGlnIuAlaIeuSer	100
DB	384	CCAGCAGCTCAACAGCTTCCACAGAAATTTTCAAGAAATGACAAAGAGCTTTCT	443
QY	101	AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleIysProGlyLeu	120
DB	444	GCCAGTGAAGTCTCCCAAGTACCTCTCAGCAATTCACATCAATCAATCAATCAAT	503
QY	121	AlaIeuSerIeuGlyValIleGlyIeuSerPheSerIeuGlnIleuValIleGlnIleuIys	140
DB	504	GCTTAAAGCTTGAAGGCTTATTCATTTAGCTTACAGCTGTGAGACATTTAAAG	563
QY	141	ProPheIleProIleHisIleMetGluAlaHisIleAlaIleuThrIleArgLeuThrAsnIys	160
DB	564	CCATTCTTCCATTCTCATATGAGGCTCATCTTACTTATGATGATGACCAATTA	623
QY	161	ValGluPheProPheLeuValIleuLeuIleSerGlyIleHisCysIleuLeuAlaIeuVal	180
DB	624	GTAATTTCTTTTATTTAGTTCTTTGATTTCTGAGGTCATGCTGTTGATTTAGT	683
QY	181	GlnGlyValSerAspPheLeuLeuLeuGlyIysSerIeuAspIleAlaProGlyIAspMet	200
DB	684	CAAGAGCTTTCAGATTTTCTGCTTCTTGAAGAGCTTGGACATACAGGTCACATG	743
QY	201	LeuAspIysValAlaAspArgIeuSerIeuIleIysHisProGluCysSerThrMetSer	220
DB	744	CTTACCAAGTGGCAAGAGAGCTTTCTTAAATTAACATCCAGAGTCTCCACATGAGT	803
QY	221	GlyGlyIysAlaIleGlnIleuIleuAlaIysGlnIleuAspArgPheHisPheAspIleIys	240
DB	804	GGTGGAAAGCCATACAGAACTTTGGCCAAACAGAAATGATTTCAATTTGACATCAA	863
QY	241	ProProLeuHisHisAlaIysAsnCysAspPheSerPheThrGlyIeuGlnHisValThr	260
DB	864	CTTCTCCATGATCATGCTTAAATTTGATTTCTTTTACCTGACCTTCAACAGCTTACT	923
QY	261	AspIysIleIleuMetIysGluIysGluGlnGlnIleGlnIleuIleuSer	280
DB	924	GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	983
QY	281	SerAlaIleAspIleAlaIleAlaThrValGlnHisIleuAlaIysIleuValIysArg	300
DB	984	TCAGACACAGATTCCTGCTCCACAGTACAGACACATGATGATCTTGTGAAAAGA	1043
QY	301	ThrHisArgAlaIleIeuPheCysIysGlnArgAspLeuIeuProGlnAsnAlaVal	320
DB	1044	ACAATCGGGCTATTTGTTTGAACAGAGACATTTGATCTTCAAAATTAATGACATA	1103
QY	321	LeuValAlaSerGlyIleValAlaSerAsnPheThrIleArgAlaIleuGlnIleu	340
DB	1104	CTGTTTCATCTGT	1163
QY	341	ThrAsnAlaThrGlnCysThrIleuLeuCysProProAlaArgLeuCysThrAspAsnGly	360

Db	1164	ACAAAGCAACAGCGACTTGTGTGTCTCCAGACTATGACATGTAATGCG	122
Qy	361	11ehec11eA1ATPaeng1y11eg1uAglEua1gA1eG1yLeuG1y11eLeuH1sAap	380
Db	1224	ATTATGATTCATCGATGATGTGATTGAAAGACTACGCTGGCTTGGCATTTTACATGAC	128
Qy	381	11eg1uG1y11eA1g1y11eG1uP1ro1y1sC1yP1ro1eG1y1V1a1Asp11eSer1y1sG1u1a1	400
Db	1284	ATAGAAAGCATTCGGCTATGACCAAAATGTCTCTTGGAGTGAACATATCAAAAGAGTT	134
Qy	401	G1yG1uA1eSer11e1y1sV1a1P1roG1u1e1u1y1sMe1eG1u11e	414
Db	1344	CGAGAAGCTTCATTAAGTACACACATTTAAAAATGAGATA	1385
RESULT 4			
ID	AAD46856	AAD46856 strand; cDNA; 1820 BP.	
AC	AAD46856;		
XX			
DT	27-JAN-2003	(first entry)	
XX			
DE	Human glycoprotease 28472 cDNA.		
XX			
KW	Human; adenosine deaminase; seven transmembrane domain receptor; cancer;		
KW	77M; glycoprotease; immune disorder; IGA deficiency; allergy; arrhythmia;		
KW	rheumatoid arthritis; diabetes; atherosclerosis; cardiovascular disorder;		
KW	hypertension; ischemic heart disease; obesity; myocardial infarction;		
KW	endothelial cell disorder; Grave's disease; psoriasis; brain disorder;		
KW	Parkinson's disease; Alzheimer's disease; haematopoietic disorder;		
KW	cerebral oedema; metabolic disorder; liver disorder; platelet disorder;		
KW	chromosome mapping; tissue typing; gene therapy; neuroprotective;		
KW	cytostatic; anorectic; cardiant; haemostatic; gene; ss.		
OS	Homo sapiens.		
XX			
TH	Key	Location/Qualifiers	
FT	CDS	146..1390	
FT		/tag= a	
FT		/product= "Human 28472 protein"	
FT		/note= "This region is specifically claimed as SEQ ID NO: 6 in claim 1 of the specification"	
XX			
PN	WO200274960-A2.		
XX			
PD	26-SEP-2002.		
XX			
PP	08-NOV-2001; 2001WO-US051427.		
XX			
XX	08-NOV-2000; 2000US-0246768P.		
PR	08-NOV-2000; 2000US-0246772P.		
PR	15-NOV-2000; 2000US-0249185P.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
PI	Leiby KR, Kapeller-libermann R, Glucksmann M;		
XX			
DR	WPI: 2002-759898/82.		
XX	P-PSDB; AAE29234.		
XX			
PT	New 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid molecules,		
PT	useful for diagnosing and treating cancer, immune, cardiovascular,		
PT	hematopoietic, brain, pain, metabolic, liver or platelet disorders, and		
PT	in pharmacogenomics.		
XX			
PS	Claim 1; Fig 8; 178pp; English.		
XX			
CC	The present invention relates to novel 38650, 28472, 5495, 65507, 81588		
CC	or 14354 nucleic acid molecules which encode adenosine deaminase, glyco-		
CC	protease or seven transmembrane domain (77M) receptor family members.		
CC	Sequences of the invention are useful in diagnosing and treating cancer		
CC	or aberrant cellular proliferation and/or differentiation (e.g. colon or		
CC	lung cancer), immune disorders (e.g. selective IGA deficiency, rheumatoid		

CC	arthritis, diabetes, allergy'), heart or cardiovascular disorders. (e.g., hypertension, atherosclerosis, arrhythmias, ischaemic heart disease, myocardial infarction, thrombus) including endothelial cell disorders (e.g. psoriasis, Grave's disease), haematopoietic disorders, brain disorders (e.g. cerebral oedema, Parkinson's or Alzheimer's disease), pain and metabolic disorders (e.g. obesity), liver disorders or platelet disorders. They are also useful in screening assays, predictive medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical trials and pharmacogenetics) and prophylactic and therapeutic methods. The nucleic acids may also be used in chromosome mapping, tissue typing and forensic biology and as surrogate markers. Sequences of the invention are also used in gene therapy. The present sequence is human glycoprotease 28472 CDNA
CC	XX
CC	Sequence 1820 BP; 543 A; 365 C; 393 G; 518 T; 0 U; 1 Other;
CC	Alignment Scores:
CC	Pred. No.: 1,24e-208 Length: 1820
CC	Score: 2088.00 Matches: 407
CC	Percent Similarity: 99.03% Conservative: 3
CC	Best Local Similarity: 98.31% Mismatches: 4
CC	Query Match: 98.26% Indels: 0
CC	DB: 6 Gaps: 0
CC	US-10-649-273-2 (1-414) x AAD46856 (1-1820)
CC	1 MetLeuIIeLeuThrLysThrAlaGlyValPhePheLysProSerLysAlaGlyValTyr 20
CC	146 ATGCTAACTCTGACATACAGACTCAGAGAGTTTAAAAACATCAAAAGGAAGCTTAT 205
CC	21 GIuPheLysArgSerPheAsnPhelSPROGlyThrLysPheLysIleValLeu 40
CC	206 GAATTTTAAAGAGTTTAAATTTTCACTCGGAACCTATTTCTCATTAATAAGTAATTG 265
CC	41 GIyIIeGIuThrSerCyAspAspThrAlaAlaValAlaSpGIuThrGIyAsnVal 60
CC	266 GGAATTGAACACTAGTTGATATATACAGACGCTCGTGTGTGATGAATGGAAATGTG 325
CC	61 LeuGIyGluAlaIleHisSerGIuThrGIuValHisLeuLysThrGIyIleValPro 80
CC	326 TTGGAGAGAAGCAATACATCCCAAACTGAAGTTCATTTAAAAACAGGTGGATGTTCC 385
CC	81 ProAlaAlaGIuGlnLeuHisArgGIuAsnIleGlnArgIleValGIuGlnAlaLeuSer 100
CC	386 CCAGAGCTCAACACAGCTTACAGAGAAATATTCACGAATATGATACAGAGCTCTTCT 445
CC	101 AlaSerGIyValSerProSerAspLeuSerAlaIleAlaThrThrIleLysProGIyLeu 120
CC	446 GCCATGTAGAGTCTCTCCAAAGTACCTCTGACGATTTGCAACTACATTAACACAGACTT 505
CC	121 AlaLeuSerLeuGIyValGIyLeuSerPheSerLeuGlnLeuValGIyGlnLeuLys 140
CC	506 GCTTTAAGCTCGAGAGTGGCTTATCATTTAGCTTACAGCTGTGTAGACAGTTAAAAAG 565
CC	141 ProPheIleProIleHisIleMetGIuAlaHisAlaLeuThrIleArgLeuThrAsnLys 160
CC	566 CCATTCATTCCTCATTCATATGAGGCTCATGACCTTAATTAAGTGTGACCAATAAA 625
CC	161 ValGIuPheProPheLeuValLeuLeuIleSerGIyGIyHisCyLeuLeuAlaLeuVal 180
CC	626 GTAGAATTCCTCTTTTAACTCTTTGATTTCTGAGAGTACACTGTCTGTGGCATTTAGTT 685
CC	181 GlnGIyValSerAspPheLeuLeuLeuGIyLysSerLeuAspIleAlaProGIyAspMet 200
CC	686 CAAAGAGTTTCAGATTTTCTGTGCTTGTGAAGATCTTTGGACATAGACACAGGTGACATG 745
CC	201 LeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGIuCySerThrMetSer 220
CC	745 CTGTGCAACCTCCCAACCAACCTCTTTAATAAACAATCCCAAGAGCTCCACCATATAGT 805
CC	221 GIyGIyLysValAlaIleGlnHisLeuAlaLysGlnGIyAsnArgPheHisPheAspIleLys 240
CC	806 GGTTGGAAAGCCATATGGAATTTGGCAACAGAAATGATTTTGCATATCAA 865

QY 241 ProProLeuHsihsalysaAnCyaspheSerPheThrglyLeuGlnhisvalThr 260
 DB 866 CCCCCCTTCATCATCTAAATAATGATTTTCTTTACTGCACTTCAACAGTACT 925
 QY 261 AspLysIleIleleuLysGlyLysGluGluGlyIleGluLysGlyGlnIleLeuSer 280
 DB 926 GATTAATAATATGAAAAACAGAAACAGAGAAAGATTTAGAAAGGGCAAAATCTGTCT 985
 QY 281 SerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCysHisLeuValLysArg 300
 DB 986 TTAGCGACGACACTTCTGCTCCACAGTACAGACACAAATGCAATGCTATCTTGGAAAGA 1045
 QY 301 ThrHisArgAlaAlaIleLeuPheCysLysGlnArgAspLeuProGlnAsnAlaVal 320
 DB 1046 ACACATCGGGCTATCTGTTTGTAAAGCAGAGACTTGTACTCAAAATATATGCACTA 1105
 QY 321 LeuValAlaSerGlyGlyValAlaAspAsnPheTyrIleArgArgAlaLeuGluIleLeu 340
 DB 1106 CTGGTTGCATCTGGTGGTGTGCAAGTAACCTCTATATCCGACAGCTCTGGAATTTTA 1165
 QY 341 ThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLeuCysThrAspAsnGly 360
 DB 1166 ACAAACGACACACAGTGCACCTTGTGTGTCCTCCACAGACTATGCATGATATATGAC 1225
 QY 361 IleMetIleAlaIleArgAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAsp 380
 DB 1226 ATTATATATTCAGATGCAATGCTATTTGAAGACTACGTCCTGGCTGGCATTTTACATGAC 1285
 QY 381 IleGluGlyIleArgTyrGluProLysCysProLeuGlyValAlaAspIleSerLysGluVal 400
 DB 1286 ATAGAAGGATCGCTGATGAAACCAAAATGTCCTCTGAGTACATATCAAAAGAGTT 1345
 QY 401 GlyGluAlaSerIleLysValProGlnIleLysMetGluIle 414
 DB 1346 GGAGAGCTTCATTAATAAGTACCAATTAATAAGATGAATA 1387
 Db
 RESULT 5
 ID ACA60887 standard; cDNA; 1820 bp.
 AC ACA60887;
 XX
 DT 08-JUL-2003 (first entry)
 DE Human cDNA 28472 encoding a glycoprotease.
 XX
 KW Human; ss; gene; cancer; aberrant cellular proliferation;
 KW differentiation; immune disorders; heart disorder; brain disorder;
 KW cardiovascular disorder; endothelial cell disorder; pain disorder;
 KW haematopoietic disorder; blood vessel disorder; metabolic disorder;
 KW liver disorder; platelet disorder; glycoprotease.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 146..1390
 FT /tag= a
 FT /product= "Glycoprotease"
 FT /note= "This CDS is specifically claimed in claim 1"
 PN US2003009017-A1.
 XX
 PD 09-JAN-2003.
 PF 08-NOV-2001; 2001US-00012140.
 XX
 PR 08-NOV-2000; 2000US-0246768P.
 PR 08-NOV-2000; 2000US-0246772P.
 PR 15-NOV-2000; 2000US-0249185P.
 XX
 PA (LEIB/) LEIBY K R.
 PA (KAP/) KAPPELLER-LIBERMANN R.

PA (GLUC/) GLUCKSMANN M A.
 XX
 PI Leiby KR, Kapeller-Libermann R, Glucksmann MA;
 XX
 DR WPI: 2003-42888/40.
 DR P-PSDB: AB09569.
 XX
 PT New isolated 38650, 28472, 5495, 65507, 81588 and 14354 nucleic acid
 PT molecules, useful for diagnosing, treating cancer, pain, or immune,
 PT heart, endothelial cell, hematopoietic, blood vessel, brain, metabolic
 PT and liver disorders.
 XX
 PS Claim 2; Fig 8; 90pp; English.
 XX
 CC The invention relates to an isolated 38650 (encoding adenosine
 CC deaminase), 28472 (encoding a glycoprotease), 5495 (encoding a 7
 CC transmembrane receptor), 65507, 81588 or 14354 nucleic acid molecule (or
 CC a sequence which is at least 60% identical to the six nucleic acids or
 CC their open reading frames, fragments of at least 15 nucleotides,
 CC naturally occurring variants, or a DNA insert of the plasmid deposited in
 CC with the American Type Culture Collection as Accession No. not defined in
 CC the specification, which encodes the amino acid sequence). Also included
 CC are a host cell containing the nucleic acids (used to produce the
 CC proteins), the encoded proteins, an antibody that selectively binds to
 CC the polypeptide, and identifying a compound that binds to/modulates the
 CC activity of the polypeptide. The nucleic acid molecules, polypeptides and
 CC methods are useful for diagnosing, treating cancer, aberrant cellular
 CC proliferation and/or differentiation, immune disorders, heart disorders,
 CC cardiovascular disorders including endothelial cell disorders,
 CC haematopoietic disorders, blood vessel disorders, brain disorders, pain
 CC and metabolic disorders, liver disorders and platelet disorders (many
 CC examples of these disorders are given in the specification). The present
 CC sequence is the Human cDNA 28472 encoding a glycoprotease
 XX
 SQ Sequence 1820 bp; 543 A; 365 C; 393 G; 518 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 1,24e-208 Length: 1820
 Score: 2088.00 Matches: 407
 Percent Similarity: 99.03% Conservative: 3
 Best Local Similarity: 98.31% Mismatches: 4
 Query Match: 98.26% Indels: 0
 DB: Gaps: 0
 US-10-649-273-2 (1-414) x ACA60887 (1-1820)
 QY 1 MetLeuIleLeuThrLysThrAlaGlyValPhePheLysProSerLysArgLysValTyr 20
 DB 146 ATGCTAATCTTGACTTAAGACTGCAGAGATTTTAAACCATCAAAAGAAAGTTTAT 205
 QY 21 GluPheLeuArgSerPheAsnPheHisProGlyThrLeuPheLeuHisLysIleValLeu 40
 DB 206 GAATTTTAAAGAACTTTAATTTTCATCTCGAAACACATTTCTTCATTAATAATGATATG 265
 QY 41 GlyIleGluThrSerCysAspAspThrAlaAlaValValAspGluThrGlyAsnVal 60
 DB 266 GGAATTAATACTATTTGTGATGATACACAGCTCTGTGGTGAATGAAATGAGAAATG 325
 QY 61 LeuGlyGluAlaIleHisSerGlnThrGluValHisLeuLysThrGlyIleValPro 80
 DB 326 TTGGAGAGACGATACATTTCCCAACTGAAGTTCATTTAAAAACAGGTGGGATTTGCT 385
 QY 81 ProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGluAlaLeuSer 100
 DB 386 CCAACAGCTCAACAGCTTCACAGAAATATTTACAGAAATATGACAGAGCTCTTTCT 445
 QY 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleLysProGlyLeu 120
 DB 446 GCCAGTGAAGTCTTCCAGAGACCTCTCAGCAATTTGCATCAATTAACCAAGACATT 505
 QY 121 AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuLys 140
 DB 506 GCTTTAAGCCTGGAGAGGCTTATTCATTTAGCTTACAGCTGATAGACAGTAAAAAG 565

QY 141 ProPheIleProIleHisHisMetGluAlaHisAlaLeuThrIleArgLeuThrAsnLys 160
 DB 566 CCATTATCCCATTCATCATATGAGAGCTCATGCACTTACTATTAGGTGACCAATATAA 625
 QY 161 ValGluPheProPheLeuValLeuLeuIleSerGlyGlyHisGlyLeuLeuAlaLeuVal 180
 DB 626 GTAGAAATTCCTTTTATGATTTCTTGAGTCTTGAGAGTCACTGCTGTTGGCAATTAGTT 685
 QY 181 GlnGlyValSerAspPheLeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMet 200
 DB 686 CAAAGGATTCAGATTTTCTGCTTCTTGAAAGCTTTGGACATAGCAACAGGTGACATG 745
 QY 201 LeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGluCySerThrMetSer 220
 DB 746 CTTGACAAAGTGGCAAGAAAGACTTTCTTTAATAAATCATCAGAGTCTCCACATGAGT 805
 QY 221 GlyGlyLysAlaAlaIleGluHisLeuAlaLysGlnGlyAsnArgPheHisAspAspIleLys 240
 DB 806 GGTCGGAAAGCCATAGAAACATTTGGCCAAACAAAGAAATAGATTCTTCAATTTTGACATCAA 865
 QY 241 ProProLeuHisHisAlaLysAsnCyAspPheSerPheThrGlyLeuGlnHisValThr 260
 DB 866 CCTCCCTTCATCATCTATAAAATGATTTTCTTTTACTGACATTCACACAGTTACT 925
 QY 261 AspLysIleIleMetLysLysGlyLysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 280
 DB 926 GATTAATAATATATATAAACAAGAAACAAGAAAGATTTGAGAAAGGAGCAATCTGCT 985
 QY 281 SerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCyHisLysValLysArg 300
 DB 986 TTAGGACGACAGACTTCTGCTCCACAGTACAGACACAAATGGCATGCTCTTGATAAGA 1045
 QY 301 ThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaVal 320
 DB 1046 ACACATCGGGCTATCTGTTTGTGAACAGACAGACATTTTACTCTCAAAATAAAGCAGTA 1105
 QY 321 LeuValAlaSerIleGlyValAlaSerAsnPheThrIleArgArgAlaLeuGlnIleLeu 340
 DB 1106 CTGTGTGATCATCTGTGTGTGTGCAAGTAACTCTTATATCCGCAAGCTCTGGAATTTTA 1165
 QY 341 ThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeuCyThrAspAsnGly 360
 DB 1166 ACAAGCGCAACAGTGCATTTGTTGTGCTCTCTCCAGACATATGACATGATATGCGC 1225
 QY 361 IleMetIleAlaIleAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAsp 380
 DB 1226 ATTATGATTCATGAGATGGAATGGAATGGAAGACTAGCTGCTGGCATTTTACATGAC 1285
 QY 381 IleGluGlyIleArgGlyGluProLysCysProLeuGlyValAlaAspIleSerLysGlyVal 400
 DB 1286 ATGAAAGGCAATCCGCTATGAAACCAAAATGCTCTCTTGAGATGACATATCAAAAGAGTT 1345
 QY 401 GlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 414
 DB 1346 GGAAGAGCTTCATTAAGTACCAATTAATAAATGAGATTA 1387
 RESULT 6
 ABB57020 standard; cDNA; 1821 BP.
 ID ABB57020;
 AC ABB57020;
 DT 30-JAN-2003 (first entry)
 XX CDNA encoding novel human glycoprotease 28472.
 DE Cancer; aberrant cell proliferation; aberrant cell differentiation;
 KW breast cancer; ovarian cancer; prostate cancer; colon cancer;
 KW lung cancer; immune disorder; heart disorder; cardiovascular disorder;
 KW endocervical disorder; hematopoietic disorder; blood vessel disorder;
 KW brain disorder; pain; metabolic disorder; liver disorder; diabetes;
 KW platelet disorder; carcinoma; sarcoma; leukemia; Hodgkin's disease;

KW autoimmune disorder; hypertension; atherosclerosis; heart failure;
 KW myocardial infarction; ischaemic heart disease; Crohn's disease;
 KW Grave's disease; Kawasaki syndrome; Raynaud's disease; aneurysm;
 KW cerebral ischaemia; peripheral neuropathy; Alzheimer's disease;
 KW Parkinson's disease; anorexia nervosa; cachexia; glycoprotease; gene; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 147..1391
 FT /tag=a
 FT /product="Glycoprotease 28472"
 FT /note="Specifically claimed in claim 1"
 FT WO200272233-A2.
 PD 03-OCT-2002.
 XX 08-NOV-2001; 2001WO-US046724.
 PE 08-NOV-2000; 2000US-0246768P.
 PR 08-NOV-2000; 2000US-0246772P.
 PR 15-NOV-2000; 2000US-0249185P.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA Leiby KR, Kapeller-Libermann R, Glucksmann M;
 PI WPI: 2003-029938/02.
 DR P-PSDB; ABG71162.
 XX New adenosine deaminase, glycoprotease and seven transmembrane domain
 PT nucleic acids and polypeptides, designated 38650, 28472, 5495, 65507,
 PT 81588 and 14354, useful for treating e.g. leukemias, Hodgkin's disease or
 PT hypertension.
 PS Claim 2; Fig 8A-B; 178pp; English.
 XX The invention describes isolated 38650, 28472, 5495, 65507, 81588 and
 CC 14354 nucleic acid molecules (I) and their encoded polypeptides (II). The
 CC 38650 nucleic acid molecule comprises a sequence encoding adenosine
 CC deaminase. The 28472 nucleic acid molecule comprises a sequence encoding
 CC a human glycoprotease. The 5495, 65507, 81588 and 14354 comprise
 CC sequences that encode a human seven transmembrane domain (7TM). The
 CC 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid and polypeptide
 CC sequences are useful for diagnosing, preventing or treating a subject
 CC with or at risk of developing a disorder, e.g. cancer or aberrant
 CC cellular proliferation and/or differentiation (e.g. breast, ovarian,
 CC prostate, colon or lung cancer), immune disorders, heart disorders,
 CC cardiovascular disorders, endocervical disorders, hematopoietic disorders,
 CC blood vessel disorders, brain disorders, pain and metabolic disorders,
 CC liver disorders or platelet disorders. These disorders include carcinoma,
 CC sarcoma, leukemia, Hodgkin's disease, autoimmune disorders,
 CC hypertension, atherosclerosis, heart failure, myocardial infarction,
 CC ischaemic heart disease, Crohn's disease, Grave's disease, Kawasaki
 CC syndrome, Raynaud's disease, aneurysm, cerebral ischaemia, peripheral
 CC neuropathy, Alzheimer's disease, Parkinson's disease, anorexia nervosa,
 CC cachexia or diabetes. This sequence encodes the novel human glycoprotease
 CC 28472
 XX
 SQ Sequence 1821 BP; 543 A; 365 C; 394 G; 518 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 1.24e-208 Length: 1821
 Score: 2088.00 Matches: 407
 Percent Similarity: 99.03% Conservative: 3
 Best Local Similarity: 98.31% Mismatches: 4
 Query Match: 98.26% Indels: 0
 ID: 10 C-ips: 5
 US-10-649-273-2 (1-414) x ABB57020 (1-1821)
 QY 1 MetLeuIleLeuThrIleLysThrAlaGlyValAlaPhePheLysProSerLysArgLysValTyr 20

```

Db      147 AGCTAAATCTTGAAGCTGACGAGAGGTTTTTTTAAACATCAAAAGAAAGTTAT 206
Qy      21 GIupheLeuArgSerPheAsnPheH;SPROGLYThrLeuPheLeuH;SLYSILLeuValLeu 40
Db      207 GAATTTTAAAGAGTTTAAATTTTCACTCTGGAACCACTATTTCTTCAATAAATGATATG 266
Qy      41 GIYILEGLuThrSerCysAspAspThrAlaIalaValAlaSPGIuThrGlyAsnVal 60
Db      267 GGAATTAAGAACTAGTGTGATGATACAGAGAGCTGCTGTGTGATGAAATCGAGAAATGTC 326
Qy      61 LeuGLYGLuAlaIleH;SSerGlnThrGluValHISLeuLysThrGlyYILEValPro 80
Db      327 TTGGGAGAGCAATACATTCCTCCAACTGAAGTTCATTTAAAAACAGGTGGATGTTCTCT 386
Qy      81 ProAlaIaIaGlnGlnLeuH;SArgGluAsnIleGLuArgIleValGlnGlnAlaLeuSer 100
Db      387 CCAAGAGCTCAACAGCTTCAAGAGAAATATTCACGAATAGTACAGAGAGCTCTTTCT 446
Qy      101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleLysProGlyLeu 120
Db      447 GCGAGTGGAGTCTCTCCAAAGTACCTCTCAGCAATGGCACTACCAATAAACCAGAGCTT 506
Qy      121 AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuLysLys 140
Db      507 GCTTTAAGCTGGAGAGGCTTATTCATTTAGCTTACAGCTGGTAGGACAGTTAAAAAG 566
Qy      141 ProPheIleProIleH;SHISMetGluAlaHISAlaLeuThrIleArgLeuThrAsnLys 160
Db      567 CCAATTCATTCCTCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 626
Qy      161 ValGluPheProPheLeuValLeuLeuIleSerGlyYHISLysLeuLeuAlaLeuVal 180
Db      627 GTAGAATTTCTTTTATGTTCTTTGATTTCTGGAGTACCTGCTGTTGGCATTTAGTT 686
Qy      181 GlnGlyValSerAspPheLeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMet 200
Db      687 CAAGAGATTTCAGATTTTCTGCTTCTTGAGAACTCTTGACATAGACACAGAGGATG 746
Qy      201 LeuAspLysValAlaIaArgArgLeuSerLeuIleLysHISProGlyCysSerThrMetSer 220
Db      747 CTTCAGACAGGTGGCAAGAGACTTCTTAAATAAATCATCAGAGTCTCCACCATAGT 806
Qy      221 GLYGLYValAlaIleGlnHISLeuAlaLysGlnGlyAsnArgPheH;SPheAspIleLys 240
Db      807 GGTGGAAAGCCATTAACATTTGGCCAAACAGAAATGATTTCACTTTGACATCAAA 866
Qy      241 ProProLeuH;SHISAlaLysAsnCysAspPheSerPheThrGlyLeuGlnHISValThr 260
Db      867 CCTCCCTTCATCATCTCAAAAATGTGATTTTCTTTTACGTGACTTCAACACGTTACT 926
Qy      261 AspLysIleIleMetLysLysGlyLysGluGluGlyIleGluLysGlyGlnIleLeuSer 280
Db      927 GATTAATAATATGAAACAGAAACAAAGAGAGTATTAGAAAGGCGCAAACTCTCTCT 986
Qy      281 SerAlaIaAspIleAlaIaThrValGlnHISThrMetAlaCysHISLeuValLysArg 300
Db      987 TCAAGACAGACATTCCTGCAAGTACACGACACACAAATGCAATGCTCTTGAAAAAGA 1046
Qy      301 ThrHISArgAlaIleLeuPheCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaVal 320
Db      1047 ACAACATCGGCTATTTCTGTTTCTTAACAGAGAGACTTTACTCTCAAAATATAGCAGTA 1106
Qy      321 LeuValAlaISerGlyValAlaIaSerAsnPheTyrlleArgArgAlaLeuGlnIleLeu 340
Db      1107 CTGGTTGCACTCGTGTGTGTCGCAAGTAACTTCTATTCGCGAGAGCTCTGAAAAATTTTA 1166
Qy      341 ThrAsnAlaThrGlnCysThrLeuLeuCysProProArgArgLeuCysThrAspAsnGly 360
Db      1167 ACAAAGCAACACAGAGCACTTTGTGTCTCTCCAGACTATTCACATGATTAAGTGC 1226
Qy      361 IleMetIleAlaIaThrAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuH;ASP 380

```

```

Db      1227 ATTAGTATTCAGTGAATGATTTGAAAGACTACGCTGCTGGCATTTCATCAGAC 1286
Qy      381 IleGluGlyIleArgTyrgluProLysCysProLeuGlyValaSPHISerLysGluVal 400
Db      1287 ATGAAAGCATTCGCTATGAAACCAAAATGTCCTTGTGAGTAGACATATCAAAAGAACTT 1346
Qy      401 GLYGLuAlaSerIleLysValProGlnLeuLysMetGluIle 414
Db      1347 GGAAGACCTTCATTAAGTACCAATTAATAATGAGAGATA 1388

RESULT 7
ADA52832
ID ADA52832 standard; cDNA; 2208 BP.
XX
XX ADA52832;
AC
XX 20-NOV-2003 (first entry)
DT
XX
XX Human coding sequence, SEQ ID 400.
DE
XX
XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KM inflammatory disease; osteoporosis; neurological disease; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX BP1293569-A2.
PN
XX
XX 19-MAR-2003.
PD
XX
XX 21-MAR-2002; 2002EP-0006586.
PF
XX
XX 14-SEP-2001; 2001JP-00328381.
PR
XX 24-JAN-2002; 2002US-0350435P.
PR
XX
XX (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
PA
XX
XX Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto U, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
XX
XX WPI: 2003-395539/38.
DR P-PSDB; ADA54471.
XX
XX
XX New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
XX Claim 1; SEQ ID NO 400; 205pp; English.
XX
XX The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
XX Sequence 2208 BP; 660 A; 454 C; 451 G; 643 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 2,15e-193 Length: 2208
Score: 1944.00 Matches: 386
Percent Similarity: 93.24% Conservative: 0
Best Local Similarity: 93.24% Mismatches: 4
Query Match: 91.48% Indels: 24
DB: 10 Gaps: 1

US-10-649-273-2 (1-414) x ADA52832 (1-2208)
Qy      1 MetLeuIleLeuThrIleLysThrAlaGlyValaPhePheLysProSerLysArgLysValTyrr 20
Db      344 ATGCTAAATCTTGACTGAAGACTGACGAGAGTTTTTTTTTTAAACCATCAAAAGAAAGTTAT 403

```

```

OY 21 GIUPELEUARGSERPHEASNPHEHISPROGLYTHREUPHEUHSISLYSILEVALEU 40
DB 404 GAATTTTTTAAAGAGTTTATTTTCAATCCGGAACACTATTTCTTCAATAAATAAGTATG 463
OY 41 GIVILEGLUTherSerCySaSPaSPThrAlaAlaValaValaSPGIUTThrGLYAsnVal 60
DB 464 GGAAATTGAAACTGATGTGATGATACAGCAGCTGCTGTGGTGGATGAAACTGGAAATG 523
OY 61 LEUGLYGLUAlaIleHisSerGlnThrGLUValaHisLeuDYSThrGLYIleValaPro 80
DB 524 TTGGGAGAGCAATACATTTCCCAACTGAAGTTTATTAAACAGGTGGGATTTGTTCT 583
OY 81 PROAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValaGlnGluAlaLeuSer 100
DB 584 CCAGCAGCTCAACAGCTTCAACAGAAATAATTCACAGATAGACAGAAGCTCTTTCT 643
OY 101 AlaserGLYValSerProSerAspLeuSerAlaIleAlaThrThrIleLeuProGLYLeu 120
DB 644 GCACAGTGGAGTCTCTCCAAAGTGAAGCTCTCAGCAATTGCACTACCAATMAAACGAGACTT 703
OY 121 AlaleuSerLeuGLYValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuValys 140
DB 704 GCTTTAAGCTGGAGAGTGGCTTATCATTTAGCTTACAGCTGGTAGACAGTTAAAAAG 763
OY 141 PropheIleProIleHisIsmetGluAlaHisAlaLeuThrIleArgLeuThrAsnLys 160
DB 764 CCATTCATTTCCCATTCATCATATAGAGGCTCATGCACTTACTATTAAGTTGACCAATAAA 823
OY 161 ValGluPheProPheLeuValIleuLeuIleSerGLYIleHisCySLeuLeuAlaLeuVal 180
DB 824 GTAGAAATTCCTTTTATGTTCTTTGATTTCTGGAGTCACTGCTGTTGGCAATTAGTT 883
OY 181 GlnGlyValSerAspPheLeuLeuLeuGLYLeuSerLeuAspIleAlaProGLYAsnMet 200
DB 884 CAAGAGTTTCAATTTTCTGCTTTGGAAGCTTTGACATAGACCAAGGAGCATG 943
OY 201 LeuAspLysValAlaArgArgLeuSerLeuIleHisIleProGLYCySerThrMetSer 220
DB 944 CTTGACAAAGTGCAGAGAGACTTCCCTTAATAAATCACAAGTCTCCACCATAGT 1003
OY 221 GLYGLYLeuAlaIleGluHisLeuAlaLysGlnGlyAsnArgPheHisAspAspIleLys 240
DB 1004 GGAGGAGAAAGCCATAGACATTTGGCCAAACAGAAATGATTTCATTTTGAATCAAAA 1063
OY 241 PROProLeuHisIleAlaLysAsnCySaSPheSerPheThrGLYLeuGlnHisValThr 260
DB 1064 CCTCCCTTCATCATGCTAAATAATTCGATTTCTTTTACTGACATTCACACAGTTACT 1123
OY 261 AspLysIleIleMetLysLysGluLysGluGlyIleGluLysGlyGlnIleLeuSer 280
DB 1124 GATAAATAATATATGAAAGAAAGAAAGAGATGTAAGAAAGGAGCAATCTGCT 1183
OY 281 SerAlaAlaAspIleAlaIleAlaThrValGlnHisIleThrMetAlaCySHISLeuValArg 300
DB 1184 TCGGCGACAGACTTGTGTCACAGTACAGACACATGGCATTCATCTTGTAAGAAAAGA 1243
OY 301 ThrHisArgAlaIleLeuPheCySlySGlnArgAspLeuProGlnAsnAsnAlaVal 320
DB 1244 AACATCGGGCTATTCTTTTGTAGACAGAGACTTGTACTCAAAATATATGCACTA 1303
OY 321 LeuValAlaSerGLYValAlaLysAsnPheThrIleArgArgAlaLeuGlnIleLeu 340
DB 1304 CTGGTTCATCTGTGTGTGTCGCAAGTAACTTCTGTATCCGCAAGCTCTGGAATTTTA 1363
OY 341 ThrAsnAlaThrGlnCySThrIleLeuCySProProProArgLeuCySThrAspAsnGly 360
DB 1364 ACNAAAGCAACACAGTCACTTGTGTGTGTCCTCCAGACTATGACATGATTAATGGC 1423
OY 361 IleMetIleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAsp 380
DB 1424 ATTATGATTTCA----- 1435
OY 381 IleGluGlyIleArgTyrgIuProLysCySProLeuGlyValaAspIleSerLysGluVal 400

```

```

DB 1436 -----TGATGTCCTTGGAGTAGACATATCAAAAGAGT 1471
OY 401 GIVILEGLUAlaserIleLysValProGlnLeuLysMetGluIle 414
DB 1472 GGAGAGCTTCCATTAAGTACCAACAATTAAATGAGATTA 1513

RESULT 8
ADQ24627
ID ADQ24627 standard; DNA; 2890 BP.
XX
AC ADQ24627;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 7447.
XX
KM soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KM ds.
XX
OS Homo sapiens.
XX
PN W02004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnick A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 7447; 210bp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual,
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cyostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 2890 BP; 869 A; 609 C; 611 G; 789 T; 0 U; 12 Other;

Alignment Scores:
Pred. No.: 3-2e-193 Length: 2890
Score: 1944.00 Matches: 386
Percent Similarity: 93.24% Conservative: 0
Best Local Similarity: 93.24% Mismatches: 4
Query Match: 91.48% Indels: 24
DB: 12 Gaps: 1

US-10-649-273-2 (1-414) x ADQ24627 (1-2890)
OY 1 MetLeuIleLeuThrLysThrAlaGlyValaPhePheLysProSerLysArgLysValTyr 20
DB 1001 ATGCTAATCTTGACATTAAGACTGCAAGAGCTTTTATTAACATCAAAAAGAAAGATTAT 1060

```


QY	21	GIupheLeuAqSgePheAsnPhetiSproGlyThrLeuPheLysGIIeValLeu	40
Db	1061	GAATTTTAAAGATTTAATTTTCACTCTGGACACTATTTCTTCATAAATAGTATG	1120
QY	41	GIyIIeGluThrSerCyaspAspThrIaaIaaIaValAspGluThrGlyAsnVal	60
Db	1121	GGAATTGAAACTAGTTGTGATGATACACAGCTGCTGTGTGTGATGAAACTGGAAATCTGT	1180
QY	61	LeuGIyGIuAaIIeHisSerGIuThrGIuValHisLeuLyGThrGlyGlyIIeValPro	80
Db	1181	TTGGAGAGAAGCAATTCATTTCCCAACTAAAGTTCATTTAAAAACGGGGGAGTTGTCTCT	1240
QY	81	ProAlaIaGIuGlnLysHisSArgGIuAsnTIEGIIaGIIeValGInGluAaLeuSer	100
Db	1241	CCAGCACTCCACAGCTTCCACAGAGAAATATTTCAACGAATAGTACAGAAAGCTCTTCTT	1300
QY	101	AlaSerGIyValSerProSerAspLeuSerAlaIaIaThrThrIIeLysProGIyLeu	120
Db	1301	GCCAGTGAAGTCTCTCCAAAGTACCTCTCAGCAATTGACACTAACAACAGCACTT	1360
QY	121	AlaLeuSerLeuGIyValGIyLeuSerPheSerLeuGlnLeuValGIyGlnLeuLys	140
Db	1361	GCTTTAACTCGGGAGTGGGCTTATCATTTAGCTTACAGCTGAGAGCAAGTAAAAAG	1420
QY	141	ProPheIIeProIIeHisHisMetGluAaHisAaIaLeuThrIIeArgLeuThrAsnLys	160
Db	1421	CCATTCAATCCATTCATCATATGAGAGCTCATGACTTACTATTAGGTAAACCAATDAA	1480
QY	161	ValGIuPheProPheLeuValLeuLeuIIeSerGIyGlyHisCyLeuLeuAaLeuVal	180
Db	1481	GTAGAATTTCTTTTATTTAGTCTTTTGATTTCTGGAGGTCACTGTCTGTGGCAATGTT	1540
QY	181	GInGIyValSerAspPheLeuLeuLeuGIyLysSerLeuAspIIeAlaIaProGIyAspMet	200
Db	1541	CAAGGAGTTTACAGATTTTCTGCTCTTGGAAGCTTTTGACATATGACACAGTGACATG	1600
QY	201	LeuAspLysValAlaArgArgLeuSerLeuIIeLysHisSProGIuCySerSerThMetSer	220
Db	1601	CTTGCAAGGTGGCAAGAACTTCTTTAATAAAAAACATCCAGAGTGTCTCCACATGAGT	1660
QY	221	GIyGIyLysAlaIIeGluHisSLeuAaIaLysGInGlyAsnArgPheHisPheAspIIeLys	240
Db	1661	GCTGGGAAGCCATGAACTTTGGCCCAACAGAAATAGATTTCAATTTGACATCAAA	1720
QY	241	ProProLeuHisHisSAlaLysAsnCyaspPheSerPheThrGIyLeuGlnHisSValThr	260
Db	1721	CCTCCCTGCATCATGCTAAAAATTGTATTTCTTTTAACTGCAGCTTCAACACGTTACT	1780
QY	261	AspLysIIeIIeMetLysLysGIuLysGIuGlyIIeGluLysGIyGlnIIeLeuSer	280
Db	1781	GATAAAAATAAATATCAAAAAAGAAAAAGAAAGGATATTTGAAAGGGGCAAAATCTGTCT	1840
QY	281	SerAlaAlaAspIIeAlaAlaThrValGlnHisSThrMetAlaCyHisSLeuValLysArg	300
Db	1841	TCAGCAGACGACATTGCTGCCACGATACGACACAAATGGCATGTCACTGTGTAAAGA	1900
QY	301	ThrHisArgAlaIIeLeuPheCySylsGlnArgAspLeuLeuProGlnAsnAsnAlaVal	320
Db	1901	ACAATCGGGCATATTTCTGTTTGTATGACAGAGAGCTTTTAACTCAAAATATAGCACTA	1960
QY	321	LeuValAlaSerGIyGlyValAaIaSerAsnPheTyrIIeArgArgAlaLeuGluIIeLeu	340
Db	1961	CTGTGTGCATCTGGGTGTGTGCAAGTAACTTGTATCCCGACAGAGCTCTGGAAATTTTA	2020
QY	341	ThrAsnAlaThrGlnCyThrLeuLeuCySProProArgLeuCyThrAspAsnGly	360
Db	2021	ACAAACGACACAGTGACCTTTGTTGTCTCTCTCCACATATGACATGATATAGCC	2080
QY	361	IIeMetIIeAlaArgPasnGlyIIeGluArgLeuArgAlaGlyLeuGlyIIeLeuHisSAsp	380
Db	2081	ATTATGATGCTCA-----	2092
QY	381	IIeGluGIyIIeArgTyrGIuProLysCySProLeuGIyValAspIIeSerLysGIuVal	400

[illegible]

CC haemostatic and thrombolytic, receptor/ligand, and antiinflammatory
 CC activities. The cDNA sequences of the invention are useful for expressing
 CC recombinant protein for analysis. The present sequence represents a novel
 CC human cDNA sequence of the invention, this sequence is an expressed
 CC sequence tag (EST) and was identified using subtractive hybridisation
 XX
 SQ Sequence 1416 BP; 441 A; 280 C; 273 G; 422 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5,31e-173	Length:	1416
Score:	1747.00	Matches:	340
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.71%	Mismatches:	0
Query Match:	82.21%	Indels:	0
DB:	8	Gaps:	0

US-10-649-273-2 (1-414) x ABX70950 (1-1416)

```

OY 74  LysThrGlyGlyIleValProProAlaIleGlnGlnLeuHisArgGluAsnIleGlnArg 93
    :::::
DB 280 AGAACAGGGGGAGTTGTTCTCCAGCAGCTCAACAGCTTCACAGAAATATTCAACGA 339

OY 94  IleValGlnGluAlaLeuSerAlaSerGlyValSerProSerAspLeuSerAlaIleAla 113
    |||||
DB 340 ATAGTACAAAGAACTCTTCTCCAGTGGAGTCTCTCAAGTGACTCTCCAGCAATTGCA 399

OY 114 ThrThrIleLysProGlyLeuAlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGln 133
    |||||
DB 400 ACTACATAAACACAGACCTTGCTTAAGCTGGAGCTGGCTTATCATTTAGCTTACAG 459

OY 134 LeuValGlyGlnLeuLysLysProPheIleProIleHisIleMetGluAlaIleAlaLeu 153
    |||||
DB 460 CTGGTGGAGAGTAAAGAAAGCCATTCCATTCATCATAGAGGCTCAGCACTT 519

OY 154 ThrIleArgLeuThrAsnLysValGluPheProPheLeuValLeuLeuIleSerGlyGly 173
    |||||
DB 520 ACTATTAGGTGACCAATAAGATGATTCCTTTTATTAGTCTTTGATTTCGAGGT 579

OY 174 HisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeu 193
    |||||
DB 580 CACTGCTGTGGCATTAGTTCAGAGGATTCAGATTTCCTGCTTGGAAAGCTTTTG 639

OY 194 AspIleAlaProGlyAspMetCysAspLysValAlaArgArgLeuSerLeuIleLysHis 213
    |||||
DB 640 GACATGACCCAGCTGACATGCTTGAACAAGTGGCAGAGACCTTCTTAAATAACAT 699

OY 214 ProGluCysSerThrMetSerGlyGlyLysAlaIleGlnHisLeuAlaLysGlnGlyAsn 233
    |||||
DB 700 CCAGAGTGCCTCCACCATGAGTGGTGGAAAGCCATAGACATTTGGCCAAACAGGAAT 759

OY 234 ArgPheHisPheAspIleLysProProLeuHisAlaLysAsnCysAspPheSerPhe 253
    |||||
DB 760 AGATTTCATTGTGACATCAAAACCTCCCTGCATCATCTAAATAATGATTTTCTTTT 819

OY 254 ThrGlyLeuGlnHisValThrAspLysIleIleMetLysLysGluLysGluGlyIle 273
    |||||
DB 820 ACTGGACTTAACACGTTACTGATTAATAATATGAAAAAGAAAAAGAGATATTT 879

OY 274 GluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaIleAlaIleAlaIleHisThrMet 293
    |||||
DB 880 GAGAAAGGGGAAATCTGTTCTTCAGCAGACATTCCTGCCACATGACACACAAATG 939

OY 294 AlaCysHisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeu 313
    |||||
DB 940 GCATGTCATTTTGAAAAGAACATCGGCTATTGTTGTTTGAAGCAGAGACTTG 999

OY 314 LeuProGlnAsnAlaValLeuValAlaSerGlyValAlaAlaSerAsnPheTyrIle 333
    |||||
DB 1000 TTACTCTCAAAATATATGAGTACTGTGTGCTGTGTGCGAAGTAACCTTCTATATTC 1059

OY 334 ArgArgAlaLeuGlnIleLeuThrAspAlaThrGlnCysThrLeuLeuCysProProPro 353
    |||||
DB 1060 CGCAGAGCTCTGAAATTTTAAACAAACGACACAGTGCACCTTGTGTGTCTCTCC 1119
  
```

```

OY 354 ArgLeuCysThrAspAsnGlyIleMetIleAlaIleTyrPheGlyIleGlnArgLeuArgAla 373
    |||||
DB 1120 AGACTATGCACTGATTAAGGATTAATGATTCATGAGATGATGTAATGAAAGACTAGTCT 1179

OY 374 GlyLeuGlyIleLeuHisAspIleGlnGlyIleArgTyrGlnProLysCysProLeuGly 393
    |||||
DB 1180 GGCCTGGCATTTTACATGACATGAGACATCCGCTATGAAACCAAAATGTCCTTGG 1239

OY 394 ValAspIleSerLysGluValGlyGluAlaSerIleLysValProGlnLeuLysMetGlu 413
    |||||
DB 1240 GTAGACATATCAAAAGAAAGTTGAGAGACCTTCATTAAGATACCAATTAATAATGAG 1299

OY 414 Ile 414
    |||
DB 1300 ATA 1302

RESULT 10
ABQ75508
ID ABQ75508 standard; DNA; 1572 BP.
XX
AC ABQ75508;
XX
DT 07-NOV-2002 (first entry)
DE
XX Murine sialoglycoprotease-like gene sequence SEQ ID NO:7.
XX
KW Murine; mouse; protease; calcium activated neutral protease type 5;
KW CARN5; trypsinase 4; sialoglycoprotease; enzyme; genetic disease;
KW neurological; neuropsychological; psychotic illness; transgenic animal;
KW gene; ds.
XX
OS Mus musculus.
XX
EN W0200245491-A2.
XX
PD 13-JUN-2002.
XX
PE 05-DEC-2001; 2001WO-US046405.
XX
PR 06-DEC-2000; 2000US-0251803P.
PR 06-DEC-2000; 2000US-0251820P.
PR 13-DEC-2000; 2000US-0255971P.
XX
PA (DELT-) DELTAGEN INC.
XX
PI Allen KD, Leviten MW;
XX
DR WPI; 2002-657389/70.
XX
PT Novel transgenic animal, comprising a disruption in protease target gene,
PT is useful for identifying agents that ameliorates a phenotype associated
PT with a disruption in a protease target gene.
XX
XX
PS Example 3; Fig 7; 62pp; English.
XX
XX The present invention describes a non-human transgenic animal (I)
XX comprising a disruption in a protease target gene (PG) selected from
XX calcium activated neutral protease type 5 (CAPN5) gene, trypsinase 4 gene
XX and sialoglycoprotease-like gene. Also described is a targeting construct
XX (II), comprising a first polynucleotide sequence homologous to at least a
XX first portion of PG, a second polynucleotide sequence homologous to at
XX least a second portion of PG and a selectable marker. (II) is useful for
XX producing a transgenic mouse comprising a disruption in a protease target
XX gene, by introducing (II) into a cell, introducing the cell into a
XX blastocyst, implanting the resulting blastocyst into a pseudopregnant
XX mouse, where the pseudopregnant mouse gives birth to a chimeric mouse,
XX and breeding the chimeric mouse to produce the transgenic mouse. (I) is
XX useful for identifying an agent that modulates the expression or function
XX of a protease target gene, by administering an agent to (I) and
XX determining whether the expression or function of the disrupted protease
XX target gene in (I) is modulated. (I) is also useful for testing the
XX efficacy of proposed genetic and pharmacological therapies for human
  
```


CC genetic diseases, such as neurological, neuropsychological or psychotic
 CC illness. The present sequence represents murine stialoglycoprotease-like
 CC gene sequence, which is used in an example from the present invention
 XX
 SQ Sequence 1572 BP; 459 A; 337 C; 340 G; 429 T; 0 U; 7 Other;

Alignment Scores:

Pred. No.:	3,766-152	Length:	1572
Score:	1549.00	Matches:	326
Percent Similarity:	84.50%	Conservative:	23
Best Local Similarity:	78.93%	Mismatches:	60
Query Match:	72.89%	Indels:	7
	6	Gaps:	2

US-10-649-273-2 (1-414) x ABQ75508 (1-1572)

```

OY 1 MetLeuIleLeuThrIsthrAlaGlyValPhePheLeuProSerIsthrGlyValTyr 20
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 5 ATGCTAATGTTAAAGAGACAGACAGAGAGCTATTCCCAAGCCCAAGAGTAAAGTTAT 64
OY 21 GluPheLeuArgSerPheAsnPheHisProGlyThrLeuPheLeuHisIsthrValLeu 40
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 65 GGATTTTAAAGAGATTAGCTTCACTCCAGAACTCTCTCTTGTCTAATACTGGTCTG 124
OY 41 GlyIleGluThrSerCysAspAspThrAlaAlaAlaValAlaAspGluThrGlyAsnVal 60
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 125 GGAATTGAACACAGCTGTATGACACAGAGCGGCTGTGTGTATGAATCGGGAATGTG 184
OY 61 LeuGlyGluAlaIleHisSerGlnThrGluValHisLeuIsthrGlyIleValPro 80
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 185 CTGGGGGAGACACTGCATCCCAACTCAGGTCATCTGAAAACAGTGGGATTTGTTCT 244
OY 81 ProAlaIleGlnGlnIleuHisArgGluAsnIleGlnArgIleValGlnGluAlaLeuSer 100
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 245 CCAAGTACTCAACACTTCCACAGAAATATTCACGAATAGTAGAAGAACTCTTCT 304
OY 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrIsthrIleValProGlyLeu 120
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 305 GGCTGTAGAAATACACCCCAAGCGATCTCTCAGCAATTCACATCAACCGGAGCTG 364
OY 121 AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnIleuValGlyIleLeuValys 140
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 365 GCCCTAAGTTGGGAGATTTGGCTTATCTTTCCTTACAGCTAATATCAGTTTAAARS 424
OY 141 ProPheIleProIleHisMetGluAlaHisAlaLeuThrIleArgLeuThrAspIle 160
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 425 CCAATTAT-CCGATTATCATCAGAGGCTCAGCGACTGACTATGAGCTCACCAGAAATAA 483
OY 161 ValGluPheProPheLeuValIleLeuIleSerGlyIsthrCysLeuLeuAlaLeuVal 180
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 484 GTAGATTN-CTTTTATTAGTTCTTTGATTCTGGCGGACATCGCTGGKRTTAGTC 542
OY 181 GlnGlyValSerAspPheLeuLeuGlyValysSerIleuAspIleAlaProGlyAspMet 200
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 543 CAAGGATTTCCGATTTCTGCTCTTGGGAGACTTTTGGACATAGACACAGGAGCATG 602
OY 201 LeuAspIleValAlaAlaArgArgLeuSerIleuIleValHisProGluIsthrMetSer 220
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 603 CTTGACAAAGGTGCAAGAAAGACTTTCTTAAATCAAAACATCCAGAAATGTTTACAAATGAGT 662
OY 221 GlyGlyIleValAlaIleGlnIleLeuAlaValysGlnGlyAsnArgPheHisPheAspIle 240
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 663 GGTGAAAAGCTTATGAACAGTTGGCCAAAGACGGAATATGATTCCATTTACTATCAAT 722
OY 241 ProProLeuHisIsthrAlaValysAsnCysAspPheSerPheThrGlyLeuGlnHisValThr 260
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 723 CCAACGATACAGAAATCTAAGATTCGATTTTCTTTTCAAGGAGCTTCAACATATTACT 782
OY 261 AspIsthrIleIleMetIsthrValysGlyValGluGlnGlyIleGluIsthrGlyIleLeuSer 280
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 783 GATTAAGCTAATATACACAAAGGAAAGAAAGAAAGGAGCATATGAGAGGCGCAAACTCTGTCA 842
OY 281 SerAlaIleAspIleAlaIsthrValGlnIsthrMetAlaCysHisLeuValIsthrArg 300

```

```

Db 843 TCAGCTGACACACTTGTCTGCTGCGGTACAGATGCCAACAGGTCACACTTGGCAAAAGA 902
OY 301 ThrHisArgAlaIleLeuPheCysValysGlnArgAspLeuLeuProGlnAsnAsnAlaVal 320
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 903 ACACATGCGCTATTCTGTTTGGACAGCAAGAAATTTGCTCTCTCCAGCTAACGCACTA 962
OY 321 LeuValAlaSerGlyGlyValAlaSerAsnPheThrIleArgArgAlaLeuGluIleLeu 340
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 963 TTAGTTTATCTGGAGGTGTGGCAAGTACCTGTACATCCGAAAGACATTGGAATTTCTC 1022
OY 341 ThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLeuCysThrAspAsnGly 360
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 1023 GCAATGCAACGATGACAGCGGTGTGTGTCTCACT-TCAGACCTGTGACCTGACAAATGGG 1081
OY 361 IleMetIleAlaThrPheAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAs 380
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 1082 CATATGATTCATGAAATGGAATGGAATGGAATGATGATGATGATGATGATGATGATGATG 1141
OY 380 P-IleGluGlyIleArgTyrGluProIsthrCysProLeuGlyVal---AspIleSerIsthr 399
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 1142 TGATAGAGACATTGCGTTATTAACCAATGTCTCTTTGATGATGATGATGATGATGATG 1201
OY 399 IuValGlyGluAlaSerIleValValProGlnLeu 410
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 1202 AAGTGGCAGAA---AGCTTGCCCATTAATAAAGTTA 1233
RESULT 11
ABS76639 standard; DNA; 1526 BP.
ID ABS76639
AC ABS76639;
XX
DT 11-DEC-2002 (first entry)
XX
DE DNA encoding novel human metalloprotease MPI fragment #1.
XX
KW Metalloprotease; MP-1; immune disorder; glutamate transport; cancer;
KW motor neuron disorder; amyotrophic lateral sclerosis; ALS; diabetes;
KW reproductive disorder; Klinefelter's syndrome; germinal cell aplasia;
KW genital wart; metabolic disorder; premature puberty; Kallman syndrome;
KW Cushing's syndrome; neurodegenerative disease; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; Tourette syndrome; sepsis;
KW liver disease; renal disease; immune disorder; rheumatoid arthritis;
KW acquired immunodeficiency syndrome; AIDS; pulmonary disease; pneumonia;
KW emphysema; cystic fibrosis; vascular disorder; inflammatory disorder;
KW neurological disorder; gene; ds.
XX
OS Homo sapiens.
XX
PN M0200272751-A2.
XX
PD 19-SEP-2002.
XX
PF 05-FEB-2002; 2002MO-US003353.
XX
PR 05-FEB-2001; 2001US-0266518P.
XX
PR 10-APR-2001; 2001US-0282814P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Chen J, Feder J, Nelson TC, Duclos F, Kryetsek S;
XX
XX WPI; 2002-723329/78.
XX
XX P-PSDB; ABG96487.
XX
PT New isolated nucleic acid encoding MP-1 protein, useful for preventing,
PT treating, or ameliorating diseases associated with aberrant
PT metalloproteinase activity, e.g. immune, metabolic, inflammatory and
PT neurological disorders.
XX
PS Disclosure; Page 462-463; 473pp; English.

```

The invention describes an isolated nucleic acid molecule (I) encoding a metalloprotease (MP-1). (I) is useful for preventing, treating, or ameliorating a medical condition, particularly an immune disorder, an aberrant glutamate transport or motor neuron disorder, such as amyotrophic lateral sclerosis (ALS), its juvenile form or an ALS-like condition. The compositions and methods are also useful for diagnosing, prognosticating, treating, ameliorating and/or treating disorders associated with MP-1 activity, e.g. diabetes, cancer, reproductive disorders (e.g. Klinefelter's syndrome, genital warts, or germinal cell aplasia), metabolic disorders (e.g. premature puberty, Kallman syndrome, or Cushing's syndrome), neurodegenerative diseases (Alzheimer's disease, Parkinson's disease, Huntington's disease or Tourette syndrome), liver and renal diseases and immune disorders (e.g. AIDS, rheumatoid arthritis or sepsis), pulmonary diseases (e.g. pneumonia, emphysema or cystic fibrosis) and vascular, inflammatory and neurological disorders (e.g. Alzheimer's disease or Parkinson's disease). This sequence represents a metalloprotease MP1 polynucleotide

CC Sequence 1526 BP; 484 A; 297 C; 289 G; 456 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.9e-135	Length:	1526
Score:	1385.00	Matches:	267
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.18%	Indels:	0
DB:	6	Gaps:	0

US-10-649-273-2 (1-414) x ABS76639 (1-1526)

QY 148 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnIleValGluPheProPheLeuVal 167
DB 1 ATGGAGGCTCATCATCTTACTATTAGGTTACCAATTAAGTAGAATTTCTTTTAAAGTT 60
QY 168 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValIleSerAspPheLeu 187
DB 61 CTTTGAATTTCTGAGAGTCACTGCTGTTGGCATTAGTCAAGAGTTTCAAGATTTTCTG 120
QY 188 LeuLeuGlyIleSerLeuAspIleAlaProGlyAspMetLeuAspIleValAlaArgArg 207
DB 121 CTTCCTGGAAGCTTGTGACATAGACACAGTGACATGCTTGAACAAGTGGCAAGAGA 180
QY 208 LeuSerLeuIleLeuHisProGluCysSerThrMetSerGlyGlyValIleGluHis 227
DB 181 CTTTCTTTAATAAACAATCCAGAGTGTCTCCACCAAGTGTGGAAAGCCATGAACAT 240
QY 228 LeuAlaIleGlnGlyAsnArgPheHisPheAspIleLeuSerProPheLeuHisAlaIle 247
DB 241 TTGGCCAAACAGAAATAGATTTCATTTCATCAACCTCCCTTCGATCATGCTTAA 300
QY 248 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspIleIleIleMetIleIle 267
DB 301 AATTGATTTTTCTTTTACGTGACCTCAACAGCTTACTGATAAATAATATGAAAAAG 360
QY 268 GluIleGlyGlnGlyIleGluIleGlyGlnIleLeuSerSerAlaAlaAspIleAlaIle 287
DB 361 GAAAAAGAGAGAGATATTTAGAAAGGGCAATCTGCTTCAGACAGACGATTTGCTGCC 420
QY 288 ThrValGlnHisThrMetAlaCysHisLeuValIleValIleValIleValIleValIle 307
DB 421 ACGATACAGACACAAATGCGATGTCATCTGTGAAAGAACACATCGGCGTATCTGTTT 480
QY 308 CysLeuGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyIleVal 327
DB 481 TGTAGACAGAGACATTTTACCTCAAAATATATGACAGTACGTGTTGATCTGATGCTC 540
QY 328 AlaSerAsnPheThrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr 347
DB 541 GCAAGTAACTTATATATCCGACAGCTCTGGAATTTTAAACAACGCAACAGTGCACCT 600
QY 348 LeuLeuCysProProAlaGluLeuCysThrAspAsnGlyIleMetIleAlaIleTyrAsnGly 367
DB 601 TTGTGTGTCTCTCTCCACAGACTATGCACTGATATATGCAATTATGATTCATGAAATGCT 660

QY 368 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGlnGlyIleArgTyrGlu 387
DB 661 ATTAAGAGACTACGTCGCTGGCTTGGCATTTTACATGACATAGAGGATCCGCTATGCA 720
QY 388 ProIleCysProLeuGlyValAspIleSerIleGlyValGlyGlnAlaSerIleLeuVal 407
DB 721 CCAAAATGTCCTCTTGGAGTAGACATATCAAAAGAGTTGAGAACTTCCATTAAGTA 780
QY 408 ProGlnLeuIleValMetGluIle 414
DB 781 CCACATTTAAATATGAGATA 801

RESULT 12

ADE31345/C
ID ADE31345 standard; DNA; 3358 BP.

AC ADE31345;

DT 29-JAN-2004 (first entry)

XX Human diagnostic and therapeutic polynucleotide (dthp), SEQ ID No 100.

DE diagnostic and therapeutic polynucleotide; dthp; antiarteriosclerotic;
XX antiinflammatory; cerebroprotective; antidiabetic; antidiabetic;
XX immunosuppressive; neuroprotective; nootropic; neuroleptic; tranquilizer;
XX osteopathic; antiarthritic; antirheumatic; cytostatic; hepatotropic;
XX vitruce; haemostatic; anti-HIV; antithyroid; thyromimetic;
XX dermatological; antibacterial; fungicide; antiparasitic; anticonvulsant;
XX chromolytic; anticoagulant; anorectic; vasodilator; antitumor;
XX gene therapy; protein replacement therapy; human; gene; ds.

OS Homo sapiens.

PN WO2003062376-A2.

PN 31-JUL-2003.

PD 13-JAN-2003; 2003WO-US001096.

PF 16-JAN-2002; 2002US-0349384P.

PR 17-JAN-2002; 2002US-0349413P.

PR 17-JAN-2002; 2002US-034946P.

PA (INCY-) INCYTE GENOMICS INC.

XX Jones AL, Dahl CR, Gietzen D, Chinn J, Dufour GE, Jackson JL,
XX Yu JY, Tuason O, Yap PE, Amshay SR, Dam TC, Liu TF, Gerstin EH;
XX Peralta CH, Lewis SA, Chen A, Marwaha R, Lan RV, Urashka ME;
XX Kristnam SR, Kolluru V, Panesar IS;

XX WPI; 2003-636732/60.

DR P-PSDB; ADE31156.

PT New human diagnostic and therapeutic polynucleotides and polypeptides,
PT useful for diagnosing, treating or preventing e.g. leukemia, brain
PT cancer, atherosclerosis, AIDS, thyroiditis, infections, obesity, stroke
PT or Alzheimer's.

XX Claim 1; SEQ ID NO 100; 634dp; English.

XX The invention relates to a novel isolated human diagnostic and
XX therapeutic polynucleotide (designated dthp). The novel dthp
XX polynucleotide comprises: any of 188 DNA sequences consisting of 195-7798
XX base pairs fully defined in the specification; a polynucleotide
XX comprising a naturally occurring polynucleotide sequence at least 90%
XX identical to the dthp polynucleotide; a polynucleotide complementary to
XX the dthp polynucleotide or its polynucleotide which is at least 90%
XX identical; or an RNA equivalent of any of the polynucleotides mentioned
XX above. The dthp polynucleotides have the following activities:

XX antiarteriosclerotic, antiinflammatory, cerebroprotective, antidiabetic,
XX antidiabetic, immunosuppressive, neuroprotective, nootropic, neuroleptic,
XX tranquilizer, osteopathic, antithyroidic, antirheumatic, cytostatic,
XX

hepatotropic, virucide, haemostatic, anti-HIV, antithyroid, thyromimetic, dermatological, antibacterial, fungicide, antiparasitic, anticonvulsant, thrombolytic, anticoagulant, anorectic, vasotropic, and antitumor. The novel dITHP polynucleotides polypeptide can be used in gene therapy and protein replacement therapy. The dithp polynucleotides or dITHP polypeptides are useful for diagnosing, preventing or treating diseases associated with the expression of human molecules. In particular, these diseases include cancers (e.g. adenocarcinoma, leukemia, melanoma, brain cancer, breast cancer, cervix cancer, bone cancer, liver cancer, lung cancer) or other cell proliferative disorders (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, polycythemia vera, primary thrombocytopoenia), autoimmune/inflammatory disorders (e.g. AIDS, Addison's disease, thyroiditis, Crohn's disease, Graves' disease, Hashimoto's thyroiditis, multiple sclerosis, osteoarthritis, rheumatoid arthritis, scleroderma, systemic lupus erythematosus), infections (e.g. viral, bacterial, fungal or parasitic infection), developmental disorders (e.g. Cushing's syndrome or epilepsy), endocrine disorders (e.g. chromoblasts, hypopituitarism, hypogonadism, gigantism, goiter) metabolic disorders (e.g. hypercholesterolemia, hypoglycaemia, diabetes, hyperlipidaemia, obesity), neurological disorders (e.g. ischemic cerebrovascular disease, stroke, Alzheimer's disease, Pick's disease, Huntington's disease, Parkinson's disease, Creutzfeldt-Jakob disease, anxiety, schizophrenia), gastrointestinal disorders (e.g. ulcers), transport disorders (e.g. akinesia or multidrug resistance), or connective tissue disorders (e.g. Paget's disease or rickets). This polynucleotide sequence represents one of the human dithp DNA sequences of the invention.

Sequence 3358 BP; 1105 A; 577 C; 601 G; 1075 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3..39e-94	Length:	3358
Score:	1001.50	Matches:	250
Percent Similarity:	44.82%	Conservative:	1
Best Local Similarity:	44.64%	Mismatches:	7
Query Match:	47.13%	Indels:	307
DB:	10	Gaps:	2

US-10-649-273-2 (1-414) x ADEB1345 (1-3358)

```

QY 157 LeuThAsnLysValGluPheProPheLeuValLeuLeuIleSerGIyGIyHISCyaleu 176
DB 3067 GTGACCAAAATAAGATAGATTCTTTTGTAGTCT-TTGATTTCTGAGGACACGTCTG 3009
QY 177 LeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlnGlyLysSerLeuAspIleAla 196
DB 3008 TGGGCAATGATTGTAAGAGATTTCAGATTCTGCTTCTGAAAGTC-TTGGACATAGCA 2950
QY 197 ProGIyAspMet----- 200
DB 2949 CCAAGTGACATGGGTTGACAGATTAAGATTATTTCTCCATCTTTTCTTATGT 2890
QY 200 ----- 200
DB 2889 TGTCCATTTCACTAAGTAGCAATAGTGTCTACACACATTCCTAAATATTTCTGAA 2830
QY 200 ----- 200
DB 2829 TTTTATCTTAAGTAAGTCAACAAAAATTCACATATGTGTGAGAAAAATAGAAAGCAGTA 2770
QY 200 ----- 200
DB 2769 GTACACGAATTAATTAATCTTAGCCTTTCTTAATAAAATGTAAAGAGTTTCATATCTGTA 2710
QY 200 ----- 200
DB 2709 CATTAAGGCTGAATATGTTGACAGATACAGTATATGTTTGGCAAAATATATGTATGAA 2650
QY 200 ----- 200
DB 2649 AGAAGCTGCTCTTAATACTAACATACACTCAAAAAAGGTAATAATAGAAATATATATAGA 2590
QY 200 ----- 200

```

```

DB 2589 TTAACATAGAGCAATTAAAGATGCAATGACAGAAATTAATACAACTTACTTAGACCA 2530
QY 200 ----- 200
DB 2529 CAGACAGGTCGCCCGCACGCCCTTTGTTTAGAATACTACAGAGGCTACTGCAATAT 2470
QY 200 ----- 200
DB 2469 ATGAAACTACACAAACACAGACGTGCTTCCACAGTGAATAATAGGAATATAGG 2410
QY 200 ----- 200
DB 2409 ACAAGTCTTAATTAATGACGTTTCATTAAGACAGTATATGTCAACTTCAAGCCATTTT 2350
QY 200 ----- 200
DB 2349 CCAACCAATAGAACAGAAACATAGACAGGGGACGTGATTTGGCTTTATTTGGGGTC 2290
QY 200 ----- 200
DB 2289 ATCATAGGAACAGAGTTGCTGCTTACCTGATATACAGTATAGTATATTTGCCAA 2230
QY 200 ----- 200
DB 2229 GTATAGCATGTTTATTCATTACAGGGGTTTGTGTTGTATAGTAATTTCAATTTATTT 2170
QY 201 -----Leu 201
DB 2169 CTTTGCATCTTTTCGTTTACAGATTAATTTTATGACTCTMAAAATATGTTTCTTT 2110
QY 202 AspLysValAlaArgArgLeuSerLeuIleLysHisProGIyCysSerThMetSerGIy 221
DB 2109 GATAG-GTGGCAAGAAAGACTTTCTTAATAAAGATCAGAGTCTCCACCATGAGTGT 2051
QY 222 -GIyLysAlaIleGluHisLeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysPr 241
DB 2050 GGGGAAAGCCATAGAACATTTGGCCAAACAAGAAATATGATTTTCATTTTGACATCAAAAC 1991
QY 241 OProLeuHisHisAlaLysAsnArgPheSerPheThrGlyLeuGlnHisValThAs 261
DB 1990 TCCCTTCATCATATCTTAAATATGATTTTCTTTTATCTGAACTTCAACAGTTACTGA 1931
QY 261 PLeuIleIleMetLysLysGlnGlyGluGlu-----GIyIleGluLysG 271
DB 1930 TAAATTAATTAATGAAAAAGAAAAAGAGAGATATTTCTTAATAGTAAAGTTGAACA 1871
QY 272 -----GIyIleGluLysG 276
DB 1870 GATAAATATTCCTGATTTGCTCTTAAATAATAGCTGCTCATTTCTGACAGTATTTGAAGG 1811
QY 276 IyGlnIleLeuSerSerIleAlaAspIleAlaAlaThrValGlnHisThrMetAlaCysH 296
DB 1810 GGCAAATCTCTGCTTTCAGACAGACATGCTGTCACAGTACACACAAATGGCATGTC 1751
QY 296 ILeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPheG 316
DB 1750 ATCTTGGAAGAAAGACATGCGGCTATTCGTTTGTGATGACAGAGACCTTTGTAACCTC 1691
QY 316 InAsnAsnAlaValLeuValAlaSerGIyValAlaSerAsnPheTyrlleArgArg 336
DB 1690 AAAATATGACAGTACGTGTCATCTGT-GGTGTGCGAAGATCTTCATATCCGCAAG 1632
QY 336 IaleuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 356
DB 1631 CTCTGAAATTTTAAAC-AAAGCAACACAGTGCATTTGTGTGTCTCTCCAGACTAT 1573
QY 356 YsthrAspArgGlyIleMetIleAlaThrPasnGlyIleGluArgLeuArgAlaGlyLeuG 376
DB 1572 GCACGTATTAATGGAATATATGATTCATGGAATGATGATTTGAAGACTAGCTGCTGG 1513
QY 376 IyIleLeuHisAspIleGluGlyIleArgTyrlleArgTyrlleArgTyrlleArgTyrlle 396

```

Db	1512	GCATTTTACATGACATAGAACGATCCGCTATGAACAAAAATGTCCTTGAGTAGACA	14
Qy	396	l e s e t y s g l u a l g l y t u a l a s e r t l e y s v a l p r o g l i n e u l y s m e t g u i l l e	414
Db	1452	TATCAAAAGAAAGTTGGAGAGCTTCATTAAGTACCAATTAATAAATGAGAGATA	1397
RESULT 13			
ABL24633			
ID	ABL24633	standard; DNA; 1601 BP.	
XX			
AC	ABL24633;		
XX			
DT	26-MAR-2002	(first entry)	
XX			
DE	Drosophila melanogaster genomic polynucleotide SEQ ID NO 25372.		
XX			
KW	Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical; gene; ds.		
XX			
OS	Drosophila melanogaster.		
PN	WO20011042-A2.		
XX			
XX	27-SEP-2001.		
PF	23-MAR-2001; 2001WO-US009231.		
XX			
PR	23-MAR-2000; 2000US-0191637P.		
XX			
PR	11-JUL-2000; 2000US-00614150.		
PA	(PERKE) PE CORP NY.		
PI	Venter JC, Adams M, Li PWD, Myers EW;		
XX			
XX	WPI; 2001-656860/75.		
PT			
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.		
XX			
PS	Claim 1; SEQ ID NO 25372; 21np + Sequence Listing; English.		
XX			
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-AB130511), expressed DNA sequences (AB101840-AB116175) and the encoded proteins (ABB57737-ABB12072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences		
CC			
CC			
XX			
SO	Sequence 1601 BP; 344 A; 477 C; 441 G; 339 T; 0 U; 0 Other;		
Alignment Scores:			
Qy			
	Pred. No.: 1,51e-64	Length: 1601	
	Score: 714.50	Matches: 156	
	Percent Similarity: 58.22%	Conservative: 60	
	Best Local Similarity: 42.05%	Mismatches: 140	
	Query Match: 33.62%	Indels: 15	
DB:	4	Gaps: 6	
US-10-649-273-2 (1-414) x ABL24633 (1-1601)			
Qy	39	VallleuGLYllleGLunhrserCyasAAPThrAlAlAlAlValAlapglunhrGly	58
Db	79	GTCTGGGATTCGAGACCTCTCGACGACACGGGATCGCCATCTGGACACACGGGC	138
Qy	59	AanValleuGLyGluAlaIleHISerGlnThrGluValHISleuYsThrNglyGlyle	78
Db	139	CGAGTGATTCGCCAATGTCTGGAGTGCACACAGAGTTCCACCCGCTATGAGGCATT	198

OY		79	VAlProProlAalaglInglInLeuHISArgGluLysmIleGlInArgIleValGlnGluA	98
Dd		199	ATTCGGCCAGGGGCCCAACATTTCACCAGCGCCGCATCAGTCCGGCTTACAGCGCTGC	258
OY		99	LeuSerAlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrTrhIleLysPro	118
Dd		259	ATGGAAGCCGGCCAAATTGAAGCCGGAACAATACTGACGGCATTCGGGTGACAGCGTCCC	318
OY		119	GlyLeuAlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeu	138
Dd		319	GGACTCGCCGCTGAGATTGCTGGTGCGGGCTGGCTTTGACAGGGCACCTGGCCCCGCTCG	378
OY		139	LysIleLysProPheIleProIleHisMetGlnAlaHisAlaLeuThrIleArgLeuThr	158
Dd		379	CAGAAGCCCTGCTGGCCCCCTTCACACATGAGAGCGACGGCTTGCAGGCCCGCATGAA	438
OY		159	Asn-----LysValGluPheProPheLeuValLeuLeuIleSerGlyGlnHISyaleu	176
Dd		439	CATCCGAGACAGATCGGCTACCTTCTCTGCTTGGCTGGCAGCGCGGCGCACTGTACAG	498
OY		177	LeuAlaLeuValGlnGlyValSerAspPheLeuLeuLeuGlyLysSerLeuAspIleAla	196
Dd		499	TTCGTGGTGGTAAACGGCCCGGTGCGCTTACGCTTTCGGCCAACACATGAGATGCG	558
OY		197	ProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlnLys	216
Dd		559	CCCGCGAGCGCTTTCAGAACATGATGCTGGCGAGCATACCGCTGCACATTTCGCCGAATAC	618
OY		217	SerThrmSerGlyGlyLysAlaIleGluHisLeuAlaLys---GlnGlyAsnArgPhe	235
Dd		619	CGCTTTCGAGACGAGAGACGGGCCATTTAGAGATCCCGCCAGCGTGGCAGAGATCCGCTG	678
OY		236	HisPheAspIleLysProProLeuHisIleAlaLysAsnCysAspPheSerPheThrGly	255
Dd		679	GCCTACAGAGTTCCGCTGCGCACTTGCACCGACGCGGAATGCMACTTCAGCTTGC CGGCG	738
OY		256	LeuGlnHisValThrAspLysIleIleMetLysLysGluLysGluGlnGlyIleGluLys	275
Dd		739	ATCAAGAACMACTCTTCGCGGCATTCGAGCGCGGCGAGCGGGAGAGCGGACACCTCCG	798
OY		276	GlyGlnIleLeuSerSerAlaAlaAspIleAlaIleAlaThrValGlnHisThrMetAlaCys	295
Dd		799	CAGCGAGTCATCAGCAACTACGGCGACTTTCGCCCGCGCTGCGCTGTCTGACGCGG	858
OY		296	HisLeuValLysArgTrhHisArgAlaIleLeuPheCys-----LysGlnArgAspLeu	313
Dd		859	CACCTGATGACCGGACGCGGCAATGAGTACTGCTCTGCGCGACAGGACAGCTC	918
OY		314	LeuProGlnAsnAsnAlaValLeuValAlaSerGlyValAlaAsenAsnPheTyIle	333
Dd		919	TTCGGTGAACCCCGCCACACTGTCTCATGTCGCGGGTGTGGCCAAACAGATGCCATA	978
OY		334	ArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysTrhLeuLysCysProProPro	353
Dd		979	TACGCCAACATCGAACATCTGGCCGCGAGTAGTGGCTCAGAGACTTCCGTCATCGAAG	1038
OY		354	ArgLeuCystrhAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuAla	373
Dd		1039	CGGTACTCTCGACAAACGGCGTATGATGCTGCGGACGCGCGTTCAGCACTG-----	1092
OY		374	GlyLeuGlyIleLeuHisAspIleGlu---GlyIleAlaGlyTrgIuProLysCysProLeu	392
Dd		1093	-----CTCCAGAGATTAAGAGCGCGACAGCGCGCTACGACTAGAC-----	1131
OY		393	GlyValAspIleSerLysGluValGlyGluA	403
Dd		1132	AGCATTAATTCAGGCGACGCGGGAATTCGCC	1164

PEST 14
AAH5110/C
ID AAH5110 standard; cDNA, 1385 BP.

AAH5110;
AC

XX 26-JUN-2001 (first entry)
 XX Human cDNA sequence SEQ ID NO:13144.
 DE Human; primer, detection, diagnosis; antisense therapy; gene therapy; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 PN EP1074617-A2.
 PD 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-00116126.
 PF 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saico K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX MPI; 2001-318749/34.
 DR
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 PS Claim 8; SEQ ID NO 13144; 2537pp + Sequence listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 CC
 XX Sequence 1385 BP; 439 A; 237 C; 256 G; 453 T; 0 U; 0 Other;
 SQ
 XX
 Alignment Scores:
 Pred. No.: 3.6e-63 Length: 1385
 Score: 700.50 Matches: 171
 Percent Similarity: 38.31% Conservative: 1
 Best Local Similarity: 38.08% Mismatches: 1
 Query Match: 32.96% Indels: 277
 DB: 4 Gaps: 1
 US-10-649-273-2 (1-414) x AAH15110 (1-1385)
 QY 74 LyeThrcGlyValIleValProProAlaAlaGlnGlnLeuHisArgGluAsnIleGlyArg 93
 DB 1346 AGAACAGGTGGATTGTTCTCCAGCAGCTCAACAGCTTCACAGAGAAAATTTCACAGA 1287

QY 94 IleValGlnGlnAlaLeuSerAlaSerGlyValSerProSerAspLeuSerAlaIleAla 113
 DB 1286 ATACTACAGAGAGCTCTTTCGACAGTGAAGTCTCTCAAGTAGACCTCTGACGATTGCA 1227
 QY 114 ThrThrIleLysProGlyLeuAlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGln 133
 DB 1226 ACTACCAATTAACACAGAGCTGCTTAAGCTCGGAGAGGCTTATCATTTAGCTTACAG 1167
 QY 134 LeuValGlyGlnLeuLysLysProPheIleProIleHisIsmetGluAlaHisAlaLeu 153
 DB 1166 CTGGTAGACAGCTTAAAGAACCAATTCATTCATTCATTCATTCATTCATTCATTCATTC 1107
 QY 154 ThrIleArgLeuThrAsnLysValGluPheProPheLeuValLeuLeuIleSerGly 173
 DB 1106 ACTATTAGCTTGACCAATTAAGTAGAATTCCTTTTAGTCTCTTGAATTCGAGAGT 1047
 QY 174 HisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeu 193
 DB 1046 CACTGCTGTTGGCATTAGTTCAAGGAGTTCAAGATTTCTGCTTCTTGAAAGCTTTG 987
 QY 194 AspIleAlaProGlyAspMetLeuAspLysVal----- 204
 DB 986 GACATGACACAGGTGACATGCTTGACAGGT- AATTAGAATTAAATTTCTCATCTTT 928
 QY 204 ----- 204
 DB 927 TTTGTTATGTTGCCATTTCACCTAGTAGCATAGATGATGCTACACCTTACCTTAA 868
 QY 204 ----- 204
 DB 867 TATTTCGAATTTATCTAGTAGTAAGTGAAGAAAAATTCACATATGTTGAGAAAAATAGA 808
 QY 204 ----- 204
 DB 807 AAGAGTAGTACACAAATTTATTAATTTCTAGCTTTCTTAATTAATGTTAAGAGTTTCAT 748
 QY 204 ----- 204
 DB 747 ATCTGTACATTAAGGCTGAATTAAGTTGACATACAGTTATGTTTCCAAATTAATGT 688
 QY 204 ----- 204
 DB 687 ATGTGAAGAAGCGTCTCGTAACCTAATCACTAGTCAAAAAAGGTAAATAAGAAATAT 628
 QY 204 ----- 204
 DB 627 ATATAGATTAAACATTAAGGACATTAAGATGCAATGACAGATTAATCAACAATTAAT 568
 QY 204 ----- 204
 DB 567 TACACACAGACAGGTTCCCGCACCCCTTTGTTTAAGTAATCAAGAGGTTACT 508
 QY 204 ----- 204
 DB 507 GCCATATATAGAAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 448
 QY 204 ----- 204
 DB 447 TAGAAGTATAGACAAGTCTTATTAATTAAGCTTCATTAAGCAAGTTATTTGCAACT 388
 QY 204 ----- 204
 DB 387 TCAAGCCATTTTCCAAACAATAGAGCAACAATAGACAGGCGAGTATGGCTCT 328
 QY 204 ----- 204
 DB 327 TATTGTTGGGTCAATCAATAGGAACAGGTTGTCTGCTTACCTGAATATCACTATAGTC 268
 QY 204 ----- 204
 DB 267 TATATTGCCAAGTATAGCATGTTTATTCATTCAGGGGTTTTTTGTTTGTAGTAAT 208

QY 204 ----- 204
Db 207 TTTCATTTATTCCTTCATCTTTTCGTTTCACAGTATTTATTTATGACTCTAA 148
QY 205 ----- 205
Db 147 AATATGTTCTTGTGATAGTGGCAAGAACCTTCTTTAATAAACAATCCAGAGTCTCC 88
QY 218 ThMetSerGlyGlyValAlaIleGluHisLeuAlaIleGlnGlyAsnArgPheHisPhe 237
Db 87 ACCATAGTGTGGGAAAGCCATGACATTTGGCCAAACAGAAATAGATTTCATTTT 28
QY 238 AspIleValSerProLeuHisHisAla 246
Db 27 GACATCAAACTCTCCTTGATCATGCT 1
RESULT 15
ABL24632/c
ID ABL24632 standard; DNA; 3656 BP.
AC ABL24632;
XX 26-MAR-2002 (first entry)
XX DT
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 25369.
XX DE
XX Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX OS
XX Drosophila melanogaster.
XX PN
XX MO200171042-A2.
XX PD
XX 27-SBP-2001.
XX PF
XX 23-MAR-2001; 2001MO-US009231.
XX PR
XX 23-MAR-2000; 2000US-0191637P.
XX PR
XX 11-JUL-2000; 2000US-00614150.
XX PA
XX (PEKE) PE CORP NY.
XX PI
XX Venter JC, Adams M, Li FWD, Myers EW;
XX DR
XX WPI; 2001-656860/75.
XX PT
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX interactions.
XX PS
XX Claim 1; SEQ ID NO 25369; 21bp + Sequence Listing; English.
XX PX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins (AB857737-
XX AB872072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 3656 BP; 1035 A; 829 C; 874 G; 918 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.5e-61 Length: 3656
Score: 687.50 Matches: 157
Percent Similarity: 55.64% Conservative: 60
Best Local Similarity: 40.26% Mismatches: 139
Query Match: 32.35% Indels: 34
DB: 4 Gaps: 7

US-10-649-273-2 (1-414) x ABL24632 (1-3656)
QY 39 ValLeuGlyIleGluThrSerCysAspAspPheHisAlaAlaValAspGluThrGly 58
Db 2578 GTCTTGCGGCATCGAGACTCTCTGCGAGCAGACGGGCATCGCATCGAGCAGCCGGC 2519
QY 59 AsnValLeuGlyGluAlaIleHisSerGlnThrGluValHis----- 72
Db 2518 CGAGTATGTCGAATGTCTGAGTCGACAGAGAGTTCCACACCCGGTGAAGTCTACA 2459
QY 73 -----LeuValThr-GlyGlyIleValPr 80
Db 2458 TCTCCGATGCGGTGCCCAATGATGAAGTGAATGATGACGCTATGAGGATTAATTC 2399
QY 80 oProAlaIleGlnIleuHisArgGluAsnIleGlnArgIleValGlnGluAlaLeuSe 100
Db 2398 GCCCAGGGCCCGAGACCTTACCGCGCCGCGATGACGTCCTACAGCGCGTGATGGA 2339
QY 100 rAlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleValProGlyLe 120
Db 2238 GCGCGCGCATTTGAAGCCGAGCAACAACTGACGGCCATCGCGTGACACGCGTCCGCGACT 2279
QY 120 uAlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnIleuValGlyIleuValu 140
Db 2278 GCGCGTGAATTTGCTGTGGTGGCGCTGCTTGACAGGACCTGCGCCGCTGCTGACAGA 2219
QY 140 sProPheIleProIleHisIsmetGluAlaHisAlaLeuThrIleArgLeuThrAsn-- 159
Db 2218 GCGCGTGTGCGCGTTCACACATGAGAGCGGACCGCTGACGCGCCGCGATGAAACATTC 2159
QY 160 ----LyValGluPheProPheLeuValLeuLeuIleSerGlyGlyHisCysLeuLeuAl 178
Db 2158 GGAGCAGATCGGCTACCTCTCTGCTGTGCGCAGCGCGCGCATGTCACTTGCT 2099
QY 178 AleuValGlnIleValSerAspPheLeuLeuGlyIleValSerLeuAspIleAlaProG 198
Db 2098 GATGCTTAACGCGCCCGCTGCGCTTACCTTGTGCGCAACATGATGATGATGCGCCCG 2039
QY 198 YAspMetLeuAspValAlaArgArgLeuSerLeuIleValHisProGluCysSerth 218
Db 2038 CGAGGCTTTGACAGATGCGTGGCGGACTGACGCTGACATTCGCGGAGTACCGCTT 1979
QY 218 rMetSerGlyGlyValAlaIleGluHisLeuAlaIleValHis--GlnGlyAsnArgPheHisP 237
Db 1978 GTGACAGAGAGACCGGCATTTGACATTCGCCGACGCGGCGACGACGATCCGCGCTTA 1919
QY 217 eAspIleValSerProLeuHisHisAlaIleValAsnGlyAspPheSerPheThrGlyLeuG 257
Db 1918 CGAGTTTCCGCGTCCACTTGGCCAGCAGCGAAGTCACTTCACTTCCGCGGATCA 1859
QY 257 nHisValThrAspValIleIleMetLeuValGlyGluGluGlyIleGluValGly 277
Db 1858 GAACAACTCTCTCCGCGCATTCGACGCGGCGGAGCGGAGCGGACGACCTCCGACCG 1799
QY 277 nIleLeuSerSerAlaIleAlaAspIleAlaThrValGlnHisThrMetAlaCysHisI 297
Db 1798 AGTCAATGCAACATTCGCGGACTTCTGCGCGCGCTGCTGTCGCGGACCT 1739
QY 297 uValValArgThrHisArgAlaIleLeuPheCys-----LysGlnArgAspLeuLeuPr 315
Db 1738 GATGACCGGACGAGCGCGGATTTGATGATCTCTCTGCGCGGACGACGACCTCTGCG 1679
QY 315 oGlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPheThrIleArgAr 335
Db 1678 TGACACCCCGCCACACTGTGTCATGTCGCGGCTGCGCAACAGATGACATATACGC 1619
QY 335 gAlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLe 355
Db 1618 CAACATTCGAACATTCGCGCGCTGATGATGATGATGATGATGATGATGATGATGATG 1559
QY 355 uCysThrAspAsnGlyIleMetIleAlaIleValAsnGlyIleGluArgLeuArgAlaGlyLe 375
Db 1558 CTGCTCCGACCAACGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1510

```

Qy      375 uGIYlLeuHhAspIleGlu--GIYlLeArgTYrGluProLyScYsProLeuGIYVa 394
Db      1509 -----CTGCAGGATTAAGAGGCCACAGCGCGCTACGACTACGAC-----AGCAT 1466
Qy      394 lAspIleSerLySGluValGIYGIuA1a 403
Db      1465 TGATATTCAGGCGCAGCGCGGATTCGCC 1438

```

Search completed: June 16, 2005, 20:14:57
 Job time : 636.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2005, 23:18:53 ; Search time 2812.5 Seconds

(without alignments)
913.751 Million cell updates/sec

Title: US-10-649-273-2

Perfect score: 2125

Sequence: 1 MLILTKAGVFPKSRKRVY.....DISKEVGASIKYQOLMEI 414

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6054689 segs, 3103772919 residues

Total number of hits satisfying chosen parameters: 12109378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000.

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool_p/US10649273/runat_15062005_111418_6138/app_query.fasta_1.1429
-DB=Published Applications NA -QFMT=fasta -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -UNIT=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=humand0.csi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MNTLEN=0
-MAXLEN=2000000000 -USER=US10649273@cgn_1_1.1034@runat_15062005_111418_6138
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSELOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FAPOP=6 -FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
3: /cgn2_6/prodata/1/pubpna/US05_PUBCOMB.seq.*
4: /cgn2_6/prodata/1/pubpna/US04_PUBCOMB.seq.*
5: /cgn2_6/prodata/1/pubpna/US03_PUBCOMB.seq.*
6: /cgn2_6/prodata/1/pubpna/US02_PUBCOMB.seq.*
7: /cgn2_6/prodata/1/pubpna/US01_PUBCOMB.seq.*
8: /cgn2_6/prodata/1/pubpna/US00_PUBCOMB.seq.*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
11: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
12: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
13: /cgn2_6/prodata/1/pubpna/US05_PUBCOMB.seq.*
14: /cgn2_6/prodata/1/pubpna/US04_PUBCOMB.seq.*
15: /cgn2_6/prodata/1/pubpna/US03_PUBCOMB.seq.*
16: /cgn2_6/prodata/1/pubpna/US02_PUBCOMB.seq.*
17: /cgn2_6/prodata/1/pubpna/US01_PUBCOMB.seq.*
18: /cgn2_6/prodata/1/pubpna/US00_PUBCOMB.seq.*
19: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
20: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
21: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
22: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
23: /cgn2_6/prodata/1/pubpna/US05_PUBCOMB.seq.*
24: /cgn2_6/prodata/1/pubpna/US04_PUBCOMB.seq.*
25: /cgn2_6/prodata/1/pubpna/US03_PUBCOMB.seq.*
26: /cgn2_6/prodata/1/pubpna/US02_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2125	100.0	2197	14	US-10-067-443-1
2	2125	100.0	2197	18	US-10-649-273-1
3	2125	100.0	2197	18	US-10-651-722-1
4	2125	100.0	2572	21	US-10-480-988-36
5	2090.5	98.4	1387	14	US-10-067-443-21
6	2090.5	98.4	1387	18	US-10-649-273-21
7	2090.5	98.4	1387	18	US-10-651-722-21
8	2090.5	98.4	1387	21	US-10-887-5534-1047
9	2088	98.3	1245	14	US-10-012-140-6
10	2088	98.3	1820	14	US-10-012-140-4
11	1944	91.5	2208	17	US-10-094-749-400
12	1944	91.5	2890	20	US-10-723-860-7447
13	1747	82.2	1416	17	US-10-120-988-177
14	1385	65.2	1526	14	US-10-067-443-23
15	1385	65.2	1526	18	US-10-649-273-23
16	1385	65.2	1526	18	US-10-651-722-23
17	1186.5	55.8	14364	14	US-10-067-443-20
18	1186.5	55.8	14364	18	US-10-649-273-20
19	1186.5	55.8	14364	18	US-10-651-722-20
20	682	32.1	1917	18	US-10-424-599-66417
21	578	27.2	1628	19	US-10-437-963-11249
22	531.5	25.0	4360	14	US-10-081-051-2
23	524	24.7	1146	17	US-10-282-122A-14674
24	502	23.6	1044	17	US-10-282-122A-26972
25	502	23.6	94750	18	US-10-672-787-38
26	494.5	23.3	1032	17	US-10-282-122A-31043
27	492.5	23.2	1026	9	US-09-815-242-7701
28	492.5	23.2	1026	17	US-10-282-122A-30016
29	488.5	23.0	1196	20	US-10-425-115-167432
30	484.5	22.8	1014	9	US-09-815-242-6207
31	484.5	22.8	1014	17	US-10-282-122A-20511
32	484.5	22.8	1014	17	US-10-282-122A-39301
33	483.5	22.8	1014	17	US-10-282-122A-41977
34	482.5	22.7	1020	17	US-10-282-122A-32254
35	481.5	22.7	1014	9	US-09-815-242-9686
36	475.5	22.4	1029	17	US-10-282-122A-6946
37	475.5	22.4	1029	17	US-10-282-122A-22020
38	475.5	22.4	1830121	17	US-10-329-670-1
39	475.5	22.4	1830121	20	US-10-158-865-1
40	469	22.1	1023	17	US-10-282-122A-15870
41	468	22.0	371	19	US-10-430-201-3118
42	468	22.0	371	19	US-10-430-201-3119
43	468	22.0	1014	17	US-10-282-122A-21633
44	467.5	22.0	1007	17	US-10-282-122A-19220
45	465.5	21.9	1065	17	US-10-282-122A-29584

ALIGNMENTS

RESULT 1
US-10-067-443-1
Sequence 1, Application US/10067443
Publication No. US20030082782A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE HIGHLY EXPRESSED IN
TITLE OF INVENTION: SPINAL CORD, MP-1
FILE REFERENCE: D0073 NP
CURRENT APPLICATION NUMBER: US/10/067,443
CURRENT FILING DATE: 2002-02-05
PRIOR APPLICATION NUMBER: US 60/265,519
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: US 60/282,814
PRIOR FILING DATE: 2001-04-10
NUMBER OF SEQ ID NOS: 71

SOFTWARE: PatentIn version 3.0
 SEQ ID NO 1
 LENGTH: 2197
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (231)..(1472)
 US-10-649-273-1

Alignment Scores:

Pred. No.:	2,076-247	Length:	2197
Score:	2125.00	Matches:	414
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-10-649-273-2 (1-414) x US-10-649-273-1 (1-2197)

```

QY 1 MetLeuIleuThrIysrThraIaGlyValPhePheIysProSerIysArgIysValTyr 20
DB 231 ATGCTTAATCTTGACTAAGACTGCAAGAGTTTATTTTAAACCACTCAAAAGAAAGTTAT 290
QY 21 GluPheLeuArgSerPheAsnPheHisProGlyThrLeuPheLeuHisIleValLeu 40
DB 291 GAATTTTAAAGAGTTTAAATTTTCAATCCGGAACACTATTTCTCTATAAATAGTATG 350
QY 41 GlyIleGluThrSerCysAspAspThrAlaAlaValaspGluThrGlyAsnVal 60
DB 351 GGAATTGAAACTGATGTGATGATACAGCAGCTGCTGTGTGATGAAACTGGAATGTG 410
QY 61 LeuGlyGluAlaIleHisSerGluThrGluValHisIleuIysrThrGlyIleValPro 80
DB 411 TTGGAGAAAGCAATACATTTCCCAAGTCAAGTTCAATTAACAGGTGGATTTGCTCT 470
QY 81 ProAlaIaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGluAlaLeuSer 100
DB 471 CCAGCAGCTCAACAGCTTCAACAGAAATATTCACAGATAGTACAGAAAGCTCTTTCT 530
QY 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleIysProGlyLeu 120
DB 531 GCAGAGGAGTCTCTCAAGTACCTCTCAGCAATTCACATCAATCAATCAATCAATCAAT 590
QY 121 AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuIys 140
DB 591 GCTTAAAGCTGGAGAGTGGCTTATATCATTTAGCTTACAGCTGTGAGACAGTTAAAG 650
QY 141 ProPheIleProIleHisIsmetGluAlaHisAlaLeuThrIleArgLeuThrAsnIys 160
DB 651 CCATTCATTTCCATTCATCATCATGAGAGCTCATGCACTTACTTATAGTTGACCAATAAA 710
QY 161 ValGluPheProPheLeuValLeuLeuIleSerGlyIleIysCysLeuLeuAlaLeuVal 180
DB 711 GTAGATTTCTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACT 770
QY 181 GlnGlyValSerAspPheLeuLeuLeuGlyIysSerLeuAspIleAlaProGlyAspMet 200
DB 771 CAAAGAGTTTCAAGATTTTCTGCTTCTTGAAGAGCTTTTGAACATGACACAGAGTACATG 830
QY 201 LeuAspIysValAlaArgArgLeuSerLeuIleIysHisProGluCysSerThrMetSer 220
DB 831 CTTGACAAAGTGGCAAGAGACTTTCTTTTAAATCAATCAAGAGTCTCACCATGAGT 890
QY 221 GlyGlyIysAlaIleGluHisLeuAlaIysGlnGlyAsnArgPheHisPheAspIleIys 240
DB 891 GGTGGGAAGCCATGACATTTGGCCCAACCAAGAAATGATTTCTATTGACATGAAA 950
QY 241 ProProLeuHisHisAlaIysAsnCysAspPheSerPheThrGlyLeuGlnHisValThr 260
DB 951 CCTCCCTTCATCATCTAAATAATGATGATTTTCTTTTAACTGAGACTTCAACAGCTTACT 1010
QY 261 AspIysIleIleMetIysIysGlyIysGluGluGlyIleGluIysGlyGlnIleLeuSer 280

```

```

DB 1011 GATAAATAATTAAGAAAGAAAGAGAGAGATATGAGAGGGGCAATCTCTGCT 1070
QY 281 SerAlaIaAspIleAlaIaThrValGlnHisThrMetAlaCysHisLeuValIysArg 300
DB 1071 TCAGCAGAGACATTCCTGCGCACAGTACAGACACAAATGACATGATCTTGTAAGAA 1130
QY 301 ThrHisArgAlaIleLeuPheCysGlnArgAspLeuLeuProGlnAsnAsnAlaVal 320
DB 1131 ACACATGGGGCTATTTCTTTTAAAGAGAGAGCTTTGATACCTCAAAATATGACATG 1190
QY 321 LeuValIaSerGlyGlyValAlaSerAsnPheThrIleArgArgAlaLeuGluIleLeu 340
DB 1191 CTGGTTCATCTGGTGTGTGTCGCAAGTAACCTTAATCCGACAGAGCTCTGGAATTTTA 1250
QY 341 ThrAsnIaThrGlnCysThrLeuLeuCysProProAlaArgLeuCysThrAspAsnIys 360
DB 1251 ACAACGCCAACACAGTCACTTTGTGTCTCTCCGACACTATGACATGATATATG 1310
QY 361 IleMetIleAlaThrAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAsp 380
DB 1311 ATTATGATTCATGGAATGCTATTTGAAGACTACGTGCTGCTGGCAATTTTACATGAC 1370
QY 381 IleGluGlyIleArgGlyArgProIysCysAspProLeuGlyValAspIleSerIysGluVal 400
DB 1371 ATGAGAGCAATCGCTATGAAACCAAAATGCTCTTGGAGTACATATCAAAAGACTT 1430
QY 401 GlyGluAlaSerIleIysValProGlnLeuIysMetGluIle 414
DB 1431 GGAAGAGCTTCCATTAATAAGTACCAAAATTAATAATGAGATA 1472

```

RESULT 2

US-10-649-273-1

Sequence 1, Application US/10649273
 Publication No. US20040043407A1
 GENERAL INFORMATION:
 APPLICANT: Bristol-Myers Squibb Company
 TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE, MP-1
 FILE REFERENCE: D0073 CNT
 CURRENT APPLICATION NUMBER: US/10/649, 273
 PRIOR FILING DATE: 2003-08-27
 PRIOR APPLICATION NUMBER: US 60/266, 518
 PRIOR FILING DATE: 2001-02-05
 PRIOR APPLICATION NUMBER: US 10/067, 443
 PRIOR FILING DATE: 2002-02-05
 PRIOR APPLICATION NUMBER: US 60/282, 814
 PRIOR FILING DATE: 2001-04-10
 NUMBER OF SEQ ID NOS: 71
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 1
 LENGTH: 2197
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (231)..(1472)
 US-10-649-273-1

Alignment Scores:

Pred. No.:	2,076-247	Length:	2197
Score:	2125.00	Matches:	414
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	18	Gaps:	0

US-10-649-273-2 (1-414) x US-10-649-273-1 (1-2197)

```

QY 1 MetLeuIleuThrIysrThraIaGlyValPhePheIysProSerIysArgIysValTyr 20
DB 231 ATGCTTAATCTTGACTAAGACTGCAAGAGTTTATTTTAAACCACTCAAAAGAAAGTTAT 290
QY 21 GluPheLeuArgSerPheAsnPheHisProGlyThrLeuPheLeuHisIleValLeu 40

```

DB 291 GAATTTTAAAGAGTTTAAATTTTCATCCGGAACACTATTTCCTCAATAAAATAGATG 350
QY 41 GYIILEGLUTRserCySaPaSPThrAlaAlaValAlaPepGluThrGlyAsnVal 60
DB 351 GGAATTGAAACTGATTGTGATGATACAGCAGCTGCTGTGTGATGAAATCGAAATGTG 410
QY 61 LeuGlyGluAlaIleHisSerGlnThrGluValHisLeuLysThrGlyIleValPro 80
DB 411 TTGGGAGAAAGCATATACATTCCTCAACCTGAAGTTCAATTAAACAGGTGGATGTTCT 470
QY 81 ProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGluAlaLeuSer 100
DB 471 CCAGAGCTCAACAGCTTCACAGAGAAATATTCACGAATAGTACAGAAAGCTTTCT 530
QY 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleLysProGlyLeu 120
DB 531 GCCAGTGGAGTCTCTCCAGATGACCTCTCGCAATGCACTACCAATAAACAGGACTT 590
QY 121 AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuLys 140
DB 591 GCTTTAAGCTCGGAGATGGCTTATCATTTAGCTTACAGCTGGTAGGACAGTTAAAG 650
QY 141 ProPheIleProIleHisHisMetGluAlaHisAlaLeuThrIleArgLeuThrAsnLys 160
DB 651 CCATTCATTCCTCCATTCATCATGAGGCTCATGCACTACTATTAGTTGACCAATATA 710
QY 161 ValGluPheProPheLeuValLeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuVal 180
DB 711 GTAGAATTCCTTTTATGTTTGTGATTTCTGAGAGTCACTGTGTTGGCACTTATGTT 770
QY 181 GlnGlyValSerAspPheLeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMet 200
DB 771 CAAGAGTTTCAATTTTCTGCTTCTGAAAGCTTTTGACATAGACAGGAGTACATG 830
QY 201 LeuAspLysValAlaArgArgLeuSerLeuIleLysHisArgProGlyCysSerThrMetSer 220
DB 831 CTTGACAGGTGCAAGAACATCTTTCTTATATAAACATCAGAGTGTCTCCACCATGAGT 890
QY 221 GlyGlyLysAlaIleGluHisLeuAlaLysGlnGlyAsnArgPheHisPheAspIleLys 240
DB 891 GGAGGAGAAAGCCATAGAACATTTGGCCAAACAGAAATGATTTCTTGTGACATCAA 950
QY 241 ProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGlyLeuGlnHisValThr 260
DB 951 CCGCTCTGATATGCTAAATAATGATTTTCTTTACTGACCTTCAACAGCTTACT 1010
QY 261 AspLysIleIleMetLysLysGluLysGluGlyIleGluLysGlyGlnIleLeuSer 280
DB 1011 GATTAATTAATAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTCT 1070
QY 281 SerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCysHisLeuValLysArg 300
DB 1071 TAGGACGACAGACATTTGCTGCACAGTACAGACACAAATGCGATCTCTTGAAAGA 1130
QY 301 ThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaVal 320
DB 1131 ACACATCGGCTATTCTGTTTGTGACAGAGACATTTGTAACCTCAAAATTAAGTCAGTA 1190
QY 321 LeuValAlaSerGlyValAlaLysAsnPheThrIleArgArgAlaLeuGluIleLeu 340
DB 1191 CTGTGTGATCTGTGTGTGTGCAAGTACTTCAATCCGACAGGCTCGGAAATTTTA 1250
QY 341 ThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeuLeuCysThrAspAsnGly 360
DB 1251 ACAGAGGCAACAGATGCACTTGTGTGTCTCTCTCCAGACATATGCACTGAATATGTC 1310
QY 361 IleMetIleAlaIleArgGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAsp 380
DB 1311 ATATATGATTCATGCAATGGAATGGAAGAACTAGCTGTGCTGGGCAATTTTACATGAC 1370
QY 381 IleGluGlyIleArgGlyLeuProLysCysProLeuGlyValAlaAspIleSerLysGluVal 400

DB 1371 ATAGAAGCATCCGCTATGAAACCAAAATGCTCTTGTGAGTAGACATATCAAAGAGTT 1430
QY 401 GlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 414
DB 1431 GGAAGAGCTTCATATAAAGTACACAAATTAATAATGAGATA 1472
RESULT 3
US-10-651-722-1
; Sequence 1, Application US/10651722
; Publication No. US20040048302A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE, MP-1
; FILE REFERENCE: D0073 DIV
; CURRENT APPLICATION NUMBER: US/10/651,722
; PRIORITY FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 10/067,443
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 1
; LENGTH: 2197
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (231)..(1472)
US-10-651-722-1
Alignment Scores:
Pred. No.: 2,07e-247 Length: 2197
Score: 2125.00 Matches: 414
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 18
US-10-649-273-2 (1-414) x US-10-651-722-1 (1-2197)
QY 1 MetLeuIleLeuThrLysThrAlaGlyValPhePheLysProSerLysArgLysValThr 20
DB 231 ATGCTAATCTTGACCTAAGACCTGCAAGATTTTAAACATCAAAAGAAAGCTTAT 290
QY 21 GluPheLeuArgSerPheAsnPheHisArgProGlyThrLeuPheLeuHisLysIleValLeu 40
DB 291 GAATTTTAAAGATTTTAAATTTTCACTCGAACATATTCTTCATATAATAGTATG 350
QY 41 GlyIleGluThrSerCySaPaSPThrAlaAlaValAlaPepGluThrGlyAsnVal 60
DB 351 GGAATTGAAACTGATTGTGATGATACAGCAGCTGCTGTGTGATGAAATCGAAATGTG 410
QY 61 LeuGlyGluAlaIleHisSerGlnThrGluValHisLeuLysThrGlyIleValPro 80
DB 411 TTGGGAGAAAGCATATACATTCCTCAACCTGAAGTTCAATTAAACAGGTGGATGTTCT 470
QY 81 ProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGluAlaLeuSer 100
DB 471 CCAGAGCTCAACAGCTTCACAGAGAAATATTCACGAATAGTACAGAAAGCTTTCT 530
QY 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleLysProGlyLeu 120
DB 531 GCCAGTGGAGTCTCTCCAGATGACCTCTCGCAATGCACTACCAATAAACAGGACTT 590
QY 121 AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuLys 140
DB 591 GCTTTAAGCTCGGAGATGGCTTATCATTTAGCTTACAGCTGGTAGGACAGTTAAAG 650
QY 141 ProPheIleProIleHisHisMetGluAlaHisAlaLeuThrIleArgLeuThrAsnLys 160

```

Db 651 CCATTCAATCCCATATCATATGAGGCTCATGACCTTACTTATGTTGACCAATMAA 710
Qy 161 ValGluPheProPheLeuValLeuLeuLeuSerGlyYhiCysLeuLeuAlaLeuVal 180
Db 711 GAGAAATTCCTTTTATGTTAGTTCTTTGATTTCTGAGGTACGTCTGTTGGCATTTAGTT 770
Qy 181 GlnGlyValSerAspPheLeuLeuLeuGlyYLeuSerLeuAspIleAlaProGlyAspMet 200
Db 771 CAAGAGGTTTCAGATTTTCTGCTTCTTGAAAGCTTTTGACATAGACCAAGGTGACATG 830
Qy 201 LeuAspLeuValAlaAlaArgArgLeuSerLeuIleYhiAspProGluCysSerThrMetSer 220
Db 831 CTGTGACCAAGGTGCAAGAAAGACTTTCTTTAATAAACAATCAAGTCTCCACCATAGT 890
Qy 221 GlyGlyYLeuAlaIleGlnHisLeuAlaYLeuGlnGlyYAsnArgPheHisPheAspIleYs 240
Db 891 GGTGGAAAGCCATAGAACATTTGGCCCAACAGAAATAGATTTCATTTTGACATCAA 950
Qy 241 ProProLeuHisHisAlaYAsnCyAspPheSerPheThrGlyLeuGlnHisValThr 260
Db 951 CTTCCCTTCATCATGCTAAATAATTTGATTTTCTTTTACTGAGACTTCAACAGCTTACT 1010
Qy 261 AspYsIleIleMetYsGlyGlyYLeuGlyGlyYIleGluYsGlyGlnIleLeuSer 280
Db 1011 GATTAATAATATATGAAAAAGAAAAAGAAAGGTAATTGAGAAAGGGCAATCTGCTCT 1070
Qy 281 SerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCysHisIleValYsArg 300
Db 1071 TCAGACGACGACATTCCTCCGACAGACACACACACAAATGGCAATCTTGTGAAAAAG 1130
Qy 301 ThrHisArgAlaIleLeuPheCysYsGlnArgAspLeuLeuProGlnAsnAlaVal 320
Db 1131 ACACATCGGCGTATTTCTTTTGTATGACGAGACAGACTTGTACTCTCAATATATGCGTA 1190
Qy 321 LeuValAlaSerGlyYValAlaAlaSerAsnPheTyIleArgArgAlaLeuGlnIleLeu 340
Db 1191 CTGGTTCATCTGTGTGTGTGTGCGCAAGTAATCTTCAATTCGCAAGGCTCGAAATTTTA 1250
Qy 341 ThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLeuCysThrAspAsnGly 360
Db 1251 ACAAACGCAACACAGGCACTTGTGTGTGTCTCCCTCCGACACTATGCACTGATTAAGGC 1310
Qy 361 IleMetIleAlaThrAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAsp 380
Db 1311 ATTAAGATTCGATCGAATGATATTGAAAGCTACGCTGCTGCTGGCATTTTACATGAC 1370
Qy 381 IleGlnGlyIleArgTyIleArgProYsCysProLeuGlyValAspIleSerIleGluVal 400
Db 1371 ATAGAAGGATCCGCTATGAAACCAAAATGCTCTTGGAAGTACATCAAAAAGAACTT 1430
Qy 401 GlyGluAlaSerIleYsValProGlnLeuYsMetGluIle 414
Db 1431 GGAGAGCTTCCATATAAAGTACCACAATTAATAATGAGATA 1472

```

RESULT 4 US-10-480-988-36

```

; Sequence 36, Application US/10480988
; Publication No. US20050069877A1
; GENERAL INFORMATION:
; APPLICANT: GANDHI, Ameena R.; KABLE, Amy E.;
; APPLICANT: SWARNAKAR, Anita; HAPALIA, April J.A.;
; APPLICANT: TRAN, Bao; DUGGAN, Brendan M.;
; APPLICANT: WARREN, Bridget A.; ISON, Craig H.;
; APPLICANT: HONCHIEL, Cynthia D.; NGUYEN, Daniel B.;
; APPLICANT: LU, Dying Aina M.; LEE, Ernestine A.;
; APPLICANT: YUE, Henry; FORSYTHE, Ian J.;
; APPLICANT: BARBOSO, Ines; RAMKUMAR, Jayalaxmi;
; APPLICANT: GRIFFIN, Jennifer A.; LI, Jona X.;
; APPLICANT: YANG, Junming; THANGAVELOU, Kavilcha;
; APPLICANT: GIETZEN, Kimberly J.; DING, Li;
; APPLICANT: BAUGHN, Mariah R.; BOROMSKY, Mark L.;
; APPLICANT: YAO, Monique G.; CHAWLA, Narinder K.;
; APPLICANT: MASON, Patricia M.; GURURAJAN, Rajagopal;

```

```

; APPLICANT: LEE, Sally; BECHA, Shanya D.;
; APPLICANT: LEE, Soo Yeun; TRAN, Uyen K.;
; APPLICANT: ELLIOTT, Vicki S.; LVO, Wen;
; APPLICANT: SPRAGUE, William W.; TANG, Y. Tom;
; APPLICANT: LU, Yan; ZEBARUDIAN, Yeganeh
; TITLE OF INVENTION: PROTEIN MODIFICATION AND MAINTENANCE MOLECULES
; FILE REFERENCE: PF-1040 USN
; CURRENT APPLICATION NUMBER: US/10/480,988
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: PCT/US02/19360
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/300,508
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/303,445
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US 60/305,405
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/311,442
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 60/314,821
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/315,992
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: US 60/378,205
; PRIOR FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 2572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7632424CBI
; US-10-480-988-36

```

Alignment Scores:

```

Pred. No.: 2,666-247 Length: 2572
Score: 2125.00 Matches: 414
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

```

US-10-649-273-2 (1-414) x US-10-480-988-36 (1-2572)

```

Qy 1 MetLeuIleLeuThrIleYsThrAlaGlyValPhePheYsProSerIleYsArgYLeuValYr 20
Db 144 ATGCTATCTTGACTAAGACGCGAGAGTTTCTTTTAAACCATCAAAAAGAAAGTTAT 203
Qy 21 GluPheLeuArgSerPheAsnPheHisProGlyThrLeuPheLeuHisIleValLeu 40
Db 204 GAATTTTAAAGAGTTTAAATTTTCAATCTCGAACACTATTTCTTCAATTAATGATATG 263
Qy 41 GlyIleGluThrIserCysAspAspThrAlaAlaValValAspGluThrGlyYAsnVal 60
Db 264 GAAATTTAAACTGTGTGATGATGATGACGACCTCTGTGTGATGAACTGGAATGTG 323
Qy 61 LeuGlyGluAlaIleHisSerGlnThrGluValHisIleuYsThrGlyYIleValPro 80
Db 324 TTGGGAAAGCAATACATTTCCCAACCTGAAGTTCATTTAAAAAACAGGTGGATTTCT 383
Qy 81 ProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGluAlaLeuSer 100
Db 384 CCAAGACCTCAACAGCTTTCACAGAGAAATATTCACAGAAATAGTAAAGAGCTTTCT 443
Qy 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleYsProGlyYLeu 120
Db 444 GCGAGTGAAGTCTTCCAAAGTGAAGCTTCAGCAATTCGAACTACATTAACCAAGGACTT 503
Qy 121 AlaLeuSerLeuGlyValGlyYLeuSerPheSerLeuGlnLeuValGlyGlnLeuYsYs 140
Db 504 GCTTTAAGCTCGGAGAGGGCTTATCTATTAGCTTACGCTGTAGAGAGCTTAATAAAG 563

```

```

QY 141 ProPhe1lePro1leHis1MetGlu1Ala1His1Ala1LeuThr1Lea1rGleuThrAsn1ys 160
Db 564 CCATTGATCCCATTCATCATATGAGGCTCATCACTTAATATAGGTTGACCAATAAA 623
QY 161 Val1Glu1Phe1Pro1Phe1Leu1Val1Leu1Leu1Ile1Ser1Gly1Yhi1Scy1Leu1Leu1Ala1Leu1Val1 180
Db 624 GTAGAAATTCCTTTTATGATTCCTTTGATTTCTTGAGAGTCACTGCTGTTGGCAATTAAGTT 683
QY 181 Glu1Gly1Val1Ser1Asp1Phe1Leu1Leu1Gly1Lys1Ser1Leu1Asp1Ile1Ala1Pro1Gly1Asp1Met 200
Db 684 CAAGAGTTTCAGATTTTCTGCTTCTTGAAAAGCTTTTGACATPACACAGGTGACATG 743
QY 201 Leu1Asp1Lys1Val1Ala1Arg1Arg1Leu1Ser1Leu1Ile1Lys1Ile1Pro1Glu1Cys1Ser1Thr1Met1Ser 220
Db 744 CTTCGACAGGTGCGCAAGACACTTCTTTTAATAAATCAATCCAGAGTGTCTCCACCATGAGT 803
QY 221 Gly1Gly1Lys1Ala1Ile1Glu1His1Leu1Ala1Lys1Gln1Lys1Asn1Arg1Phe1His1Asp1Ile1Lys 240
Db 804 GGTGGGAAAGCCATAGAACATTTTGCCAAACAGAAATAGATTTCAATTTGACATCAAA 863
QY 241 Pro1Pro1Leu1His1His1Ala1Lys1Asn1Cys1Asp1Phe1Ser1Phe1Thr1Gly1Leu1Gln1His1Val1Thr 260
Db 864 CCTCCCTTCATCATGCTTAAATAATTTGATTTTCTTTTACTGACATTCACACAGTTACT 923
QY 261 Asp1Lys1Ile1Ile1Met1Lys1Lys1Glu1Lys1Gln1Glu1Gly1Ile1Glu1Lys1Gln1Ile1Lys1Ser 280
Db 924 GATTAATAATTAATAGAAAAGAAAAGAAAAGAGAGATTTGAAAGGGGCAAAATCTGTTCT 983
QY 281 Ser1Ala1Asp1Ile1Ala1Ala1Thr1Val1Gln1His1Thr1Met1Ala1Cys1His1Leu1Val1Lys1Arg 300
Db 984 TCAGCGACACACATGCTGCTCCACAGTACAGACACAAATGCGATGTCATCTTGTAAGAA 1043
QY 301 Thr1His1Arg1Ala1Ile1Leu1Phe1Cys1Lys1Gln1Arg1Asp1Leu1Pro1Gln1Asn1Ala1Val1 320
Db 1044 ACACATGGGCTATTTCTGTTTGTAAACAGAGACTTTGTAACCTCAAAATAATGCACTA 1103
QY 321 Leu1Val1Ala1Ser1Gly1Val1Ala1Ser1Asn1Phe1Thr1Lea1rGleuThr1Leu1Val1Leu1 340
Db 1104 CTGTTGTCATCTGTTGTTGTGTCGAAAGTAACTTCAATATCCGACAGCTCTGGAATTTTAA 1163
QY 341 Thr1Asn1Ala1Thr1Gln1Cys1Thr1Leu1Leu1Cys1Pro1Pro1Arg1Leu1Cys1Thr1Asp1Asn1Gly 360
Db 1164 ACAACGCCACACAGTGCACCTTTGTTGTGCTCTCTCCACAGCTATGCACTGATATATGCT 1223
QY 361 Ile1Met1Ile1Ala1TPa1Snc1Gly1Ile1Glu1Arg1Leu1Arg1Ala1Gly1Leu1Gly1Ile1Leu1His1Asp 380
Db 1224 ATATATGATTCATGAGATGGAATTTGAAGAATTAAGTACGTCGCTGCGCATTTTACATGAC 1283
QY 381 Ile1Glu1Gly1Ile1Arg1Gly1Leu1Pro1Lys1Cys1Pro1Leu1Gly1Val1Asp1Ile1Ser1Lys1Gln1Val1 400
Db 1284 ATAGAAAGCATCCCTCTATGAAACCAAAATGTCCTCTTGAGATGACATATCAAAAGAACTT 1343
QY 401 Gly1Glu1Ala1Ser1Ile1Lys1Val1Pro1Gln1Leu1Met1Glu1Ile 414
Db 1344 GAGAAAGCTTCATTAATAAGTACCAATTAATAATAATGAGATTA 1385

```

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 1387
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-067-443-21

Alignment Scores:
Pred. No.: 1,59e-243 Length: 1387
Score: 2090.50 Matches: 412
Percent Similarity: 93.85% Conservative: 0
Best Local Similarity: 93.85% Mismatches: 2
Query Match: 98.38% Indels: 25
DB: 14 Gaps: 1

US-10-649-273-2 (1-414) x US-10-067-443-21 (1-1387)

QY 1 Met1Leu1Ile1Leu1Thr1Lys1Thr1Ala1Gly1Val1Phe1Phe1Lys1Pro1Ser1Lys1Arg1Lys1Val1Tyr 20
Db 24 ATGCTAATCTTGACATTAAGACATGCGAGAGCTTTTAAACATCAATCAAAAAGAAAGCTTTAT 83
QY 21 Glu1Phe1Leu1Arg1Ser1Phe1Asp1Phe1His1Pro1Gly1Thr1Leu1Phe1Leu1His1Lys1Ile1Val1Leu 40
Db 84 GAATTTTAAAGAGCTTTTAATTTTCATCTGAAACACTAATTTCTTCATTAATAATAGTATG 143
QY 41 Gly1Ile1Glu1Thr1Ser1Cys1Asp1Asp1Thr1Ala1Ala1Val1Val1Asp1Glu1Thr1Gly1Asn1Val1 60
Db 144 GCAATTTAAACTAGTGTGATGATATACAGACAGCTGCTGCTGCTGATGAAATCGAAATGTG 203
QY 61 Leu1Gly1Glu1Ala1Ile1His1Ser1Gln1Thr1Glu1Val1His1Leu1Lys1Thr1Gly1Gly1Ile1Val1Pro 80
Db 204 TTGGAGAAAGCAATACATCTCCAAATCTGAAGTTCAATTAATAAAGAGGTGGATGTTCT 263
QY 81 Pro1Ala1Ala1Gln1Gln1Leu1His1Arg1Lys1Ile1Gln1Arg1Ile1Val1Gln1Glu1Ala1Lys1Ser 100
Db 264 CCAGCAGCTCAACAGCTTCCACAGAAATATTTCAACCAATAGTCAAGAAAGCTTTTCT 323
QY 101 Ala1Ser1Gly1Val1Ser1Pro1Ser1Asp1Leu1Ser1Ala1Ile1Ala1Thr1Thr1Lys1Pro1Gly1Leu 120
Db 324 GCCAGTGAAGTCTCTCCAAAGTACCTCTCGAGAAATTTGCACTACATTAACACAGAGACTT 383
QY 121 Ala1Leu1Ser1Leu1Gly1Val1Gly1Leu1Ser1Phe1Ser1Leu1Gln1Leu1Val1Gly1Gln1Leu1Lys1 140
Db 384 GCTTTAACTCGGAGAGTGGCTTATCATTTAGCTTACAGCTGATGAGACAGTAAATAAAG 443
QY 141 ProPhe1lePro1leHis1MetGlu1Ala1His1Ala1LeuThr1Lea1rGleuThr1Asn1ys 160
Db 444 CCATTGATCCCATTCATCATATGAGGCTCATGCACTTAATAGGTTGACCAATAAA 503
QY 161 Val1Glu1Phe1Pro1Phe1Leu1Val1Leu1Leu1Ile1Ser1Gly1Yhi1Scy1Leu1Leu1Ala1Leu1Val1 180
Db 504 GTAGAAATTCCTTTTATGATTCCTTTGATTTCTTGAGAGTCACTGCTGTTGGCAATTAAGTT 563
QY 181 Glu1Gly1Val1Ser1Asp1Phe1Leu1Leu1Gly1Lys1Ser1Leu1Asp1Ile1Ala1Pro1Gly1Asp1Met 200
Db 564 CAAGAGTTTCAGATTTTCTGCTTCTTGAAAAGCTTTTGACATPACACAGGTGACATG 623
QY 201 Leu1Asp1Lys1Val1Ala1Arg1Arg1Leu1Ser1Leu1Ile1Lys1Ile1Pro1Glu1Cys1Ser1Thr1Met1Ser 220
Db 624 CTTCGACAGGTGCGCAAGACACTTCTTTTAATAAATCAATCCAGAGTGTCTCCACCATGAGT 683
QY 221 Gly1Gly1Lys1Ala1Ile1Glu1His1Leu1Ala1Lys1Gln1Lys1Asn1Arg1Phe1His1Asp1Ile1Lys 240
Db 684 GGTGGGAAAGCCATAGAACATTTTGCCAAACAGAAATAGATTTCAATTTGACATCAAA 743
QY 241 Pro1Pro1Leu1His1His1Ala1Lys1Asn1Cys1Asp1Phe1Ser1Phe1Thr1Gly1Leu1Gln1His1Val1Thr 260
Db 744 CCTCCCTTCATCATGCTTAAATAATTTGATTTTCTTTTACTGACATTCACACAGTTACT 803
QY 261 Asp1Lys1Ile1Ile1Met1Lys1Lys1Glu1Lys1Gln1Glu1Gly1Ile1Glu1Lys1Gln1Ile1Lys1Ser 280
Db 804 GATTAATAATTAATAGAAAAGAAAAGAAAAGAGAGATTTGAAAGGGGCAAAATCTGTTCT 983

```


[illegible]

RESULT 7

```

US-10-651-722-21
; Sequence 21, Application US/10651722
; Publication No. US20040048302A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE, MP-1
; FILE REFERENCE: D0073 DIV
; CURRENT APPLICATION NUMBER: US/10/651,722
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 10/667,443
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 1387
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-651-722-21

```

Alignment Scores:

Pred. No.:	1.59e-243	Length:	1387
Score:	2090.50	Matches:	412
Percent Similarity:	93.85%	Conservative:	0
Best Local Similarity:	93.85%	Mismatches:	2
Query Match:	98.38%	Indels:	25
DB:	18	Gaps:	1

US-10-649-273-2 (1-414) X US-10-651-722-21 (1-1387)

Oy	1	MettLeu11LeuThrThyPheThrAlaGlyValPhePheLysProSerLysArgLysValL Tyr	20
Db	24	ATGCAATCTTGACCTGACGCTGAGGAGTTTTTTTAAACCAACAAAGAGAAAGTTAT	83
Oy	21	GlupheLeuArgSerPheasnPheHisProGlyThrLeuPheLeuHisLysValLeu	40
Db	84	GAATTTTTAAGAAAGTTTTTAATTTTCATCCGAAAGCTAATTTCTTCATAAATAATGATTC	143
Oy	41	GlyI1IegLuthSerCysAspAspThrAlaAlaValValAspG1uthrG1yAsnVal	60
Db	144	GGAAATTGAACTAGTTGTGATGATACAGCAGCGTGTGTGATGAAACTGAAATATGTC	203
Oy	61	LeuGlyGluAla11ehisSerGlnThrGluValHisLeuLysThrGlyGly11LeuAlPro	80
Db	204	TTGGGAGAAGCAATACATTTCCCAACTGAAAGTTCAATTAAAAACAGGTGGATTTGTTCT	263
Oy	81	ProAlaAlaGlnGlnLeuHisArgGluAsn11IeglnArg11LeuAlngluAlaLeuSer	100
Db	264	CCAGCAGGCTCAACAGCTTCACAGAGAAATAATTCAACGATATGACAGAGGCTCTTCT	323
Oy	101	AlaSerGlyValSerProSerAspLeuSerAla11Ala1AthrThrLysProGlyLeu	120
Db	324	GCCAGTGAAGTCTCTCCAAAGTGACCTCTCAGCAATTCGAACTACCATPAAACCGAGCTT	383
Oy	121	AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuLys	140
Db	384	GCTTTAAGCGCTGGAGTGGGCTTATCATTTAAGCTTCACGCTGTAGGACAGTTAAAAAAG	443
Oy	141	ProPhe11IePro11ehisIstMetGluAlaHisAlaLeuThr11LeuArgLeuThrAsnLys	160
Db	444	CCATTCAATCCCAATTCATCATATGAGAGGCTCAAGCACTTACATATATAGTTGACCAATPAA	503

OY	161	ValGluPheProPheLeuValLeuLeuIleSerGlyGlyHisCySbleuLeuAlaLeuVal	180
Db	504	GTAAGATTTTCCTTTTAAAGTCTTTTGAATTTCTGGAGTCACTGTCGTGGCATTAAGT	563
OY	181	GlnGlyValSerAspPheLeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMet	200
Db	564	CAAGGAGTTTCAGATTTTCTGCTCTTGGAAGAAGCTTTGGACATAGCACAGTGACATG	623
OY	201	LeuAspIlyValAlaAlaArgArgLeuSerLeuIleLysHisProGluCySerThrMetSer	220
Db	624	CTTACCAAGGCGGCAAGAACCTTTCTTTAAATAAACATCCAGAGTGTCTCCACCATGAGT	683
OY	221	GlyGlyLysValAlaIleGlnHisLeuAlaLysGlnGlyAsnAcgPheHisPheAspIleLys	240
Db	684	GGTGGGAAGCCATAGAGCATTTGGCCAAACAGAAATAGATTTCATTTGACATCAAA	743
OY	241	ProProLeuHisHisAlaLysAsnCyAspPheSerPheThrGlyLeuGlnHisValThr	260
Db	744	CTCTCCCTTGATCATGCTCTAAATTTGCAATTTTCTTTTACTGCACTTCAACAGCTTACT	803
OY	261	AspIlyValIleMetLysLysGlyLysGlnGluGlyIle-----	273
Db	804	GATAAATATTAATGAAGAAAGAAAAGAGAAAGATATTTCTTAATTAGTAAAGTTCAA	863
OY	274	-----GluLys	275
Db	864	CAGATTAATATTCCTGATGTGCTCTAAATATAGCTGCTCATTTCTCGAGGTATGAGAG	923
OY	276	GlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaLys	295
Db	924	GGGCAAAATCCGTCTTTCAGCAGCAGACATTCCTGCCACAGTACAGCACACATGGCAGT	983
OY	296	HisLeuValLysAcgThrHisAspArgAlaIleLeuPheCyLysGlnArgAspLeuLeuPro	315
Db	984	CATCTTGGAAGAAAGACATCGGGCTATTTCTGTTTTGTAAAGCAGAGAGCATTTGTAAGT	1043
OY	316	GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaLaseAspPheTyrIleArgArg	335
Db	1044	CAAAATATATGCAAGTACGCTTGTCATCTGCTGCTGCGCAAGTAATCTTATATCGCAGA	1103
OY	336	AlaLeuGluIleLeuThrAsnAlaThrGlnCyThrLeuLeuCySProProProArgLeu	355
Db	1104	GCTCTGGAAATTTTAAACAAAGCAACACAGGCACTTTGTGTGTCCTCTCCAGACTTA	1163
OY	356	CysThrAspAsnGlyIleMetIleAlaThrAsnGlyIleGluArgLeuArgAlaGlyLeu	375
Db	1164	TGCACGTATATGGCATTTATATATTCATGAGATGTATTTGAAGACTACGTGGGCTTG	1222
OY	376	GlyIleLeuHisAspIleGlyGlyIleArgTyrGlnProLysCySProLeuGlyValAsp	395
Db	1224	GGCATTTTATACATGACATGAGAGGACATCGCTATGAACCAAAATGTCTCTTGAGATAGAC	1283
OY	396	IleSerLysGlyValAlaGlyGlnHisSerIleLysValProGlnLeuLysMetGlnIle	414
Db	1284	ATATCAAAAGAAAGTGGAGAGCTTCCATTAAGTACCAACATTTAAATATGAGATGA	1340

RESULT 8

```

US-10-887-553A-1047
; Sequence 1047, Application US/1088753A
; Publication No. US20050085436A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Garza, Dan
;
; APPLICANT: Li, Hao
;
; TITLE OF INVENTION: Method to treat conditions associated
; with insulin signalling dysregulation
;
; FILE REFERENCE: 4-33262
;
; CURRENT APPLICATION NUMBER: US/10/887,553A
;
; CURRENT FILING DATE: 2004-07-09
;
; PRIOR APPLICATION NUMBER: 60/485,883
;
; PRIOR FILING DATE: 2003-08-07
;
; NUMBER OF SEQ ID NOS: 1208
;
; SOFTWARE: FastSeq for Windows Version 4.0

```


SEQ ID NO 1047	LENGTH: 1387	TYPE: DNA	ORGANISM: human	US-10-887-553A-1047
Alignment Scores:				
Pred. No.:	1.59e-243	Length:	1387	
Score:	2090.50	Matches:	412	
Percent Similarity:	93.85%	Conservative:	0	
Best Local Similarity:	93.85%	Mismatches:	2	
Query Match:	98.38%	Indels:	25	
DB:	21	Gaps:	1	
US-10-649-273-2 (1-414) x US-10-887-553A-1047 (1-1387)				
QY	1 MetLeuIleLeuThrIlySthrAlaGlyValPhePheIysProSerIlySarGlyValIyr	20		
DB	24 ATGTAATCTTGACCTAGACAGCTGACGAGAGTTTATTTTAAACATCCAAAGGAAGTTAT	83		
QY	21 GluPheLeuAIGSerPheAsnPheHisProGlyThrLeuPheLeuHisIlyValLeu	40		
DB	84 GAATTTTAAAGATTTTAAATTTTCACTCGAACAACATATTTCTTCAATAAATGATATG	143		
QY	41 GlyIleGluThrSerCysASPSPThrAlaAlaIaIaValIaASPgluThrGlyAsnVal	60		
DB	144 GGAATTTGAATCTAGTGTGATGATGATACACACACTGCTGTGGATGAGAACTGGAAATGTG	203		
QY	61 LeuGlyGluAlaIleHisSerGlnThrGluValHisLeuIySThrGlyGlyIleValPro	80		
DB	204 TTGGAGAGAGCAATACATTCCTCCAAACTGAAGTTCATTTAAACAGGTGGATTTGTTCT	263		
QY	81 ProAlaIaGlnGlnLeuHisIarGluAsnIleGlnArgIleValGlnGluAlaLeuSer	100		
DB	264 CCACAGCTCAACAGCTTCCACAGAGAAATATTTCAACGAATAGTACAAAGACTCTTTCT	323		
QY	101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleIySProGlyLeu	120		
DB	324 GCCAGTGAAGCTCTCCAAAGTACCCTTCACCAATTTGCATTCACATTAACCAAGACTT	383		
QY	121 AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuIyS	140		
DB	384 GCTTAAACCTGGAGAGTGGCTTTATCATTTAGCTTACAGCTGTGAGACATTTAAAAAG	443		
QY	141 ProPheIleProIleHisIleMetGluAlaHisIaIaLeuThrIleArgLeuThrAsnIyS	160		
DB	444 CCATTCATTTCCCATTCATCATATGAGAGGCTCATGCACTTACTATTAAGTTGACCAATAA	503		
QY	161 ValGluPheProPheLeuValLeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuVal	180		
DB	504 GTAAGATTTCTTTTAAATTTAGTTCTTTTGAATTTCTGAGAGTCACTGTCTGGCATTTAGTT	563		
QY	181 GlnGlyValSerAspPheLeuLeuLeuGlyIySLeuSerLeuAspIleAlaProGlyAspMet	200		
DB	564 CAAAGAGTTTCAAGATTTTCTCTCTTGGAAAGCTTTTGGACATAGCAACAGGTGAATG	623		
QY	201 LeuAspIySValaIaArgArgLeuSerLeuIleIySHisProGlyCysSerThrMetSer	220		
DB	624 CTTGACAAAGTGGCAAGAGACTTCTTTTAATAAATTCACAGAGTGTCTCCACATGAGT	683		
QY	221 GlyIyIyValaIaIleGluHisLeuAlaIySglnIyAsnArgPheHisPheAspIleIyS	240		
DB	684 GTGGGAAAGCCATAGAGCATTTTGGCCAAACAAAGAAATATATATTTTCTTTGGACATCAA	743		
QY	241 ProProLeuHisIleAlaIyAsnCysAspPheSerPheThrGlyLeuGlnHisValThr	260		
DB	744 CTTCTCTTGACATACGCTTAATAATTTGATATTTTCTTTTACTGAGACTTCAACAGTTACT	803		
QY	261 AspIySleIleMetIyIySglnIySglnIyIle	273		
DB	804 GATTAATTAATTAATGAAGAAAGGAAAGAGAGAGATATATTTCTTAATTAGTAAGTTGAA	863		
QY	274	GlnIyS	275	

```

Db      864 CAGATAATATATTCCTGCAATGTGCTTAATAAATAGCGCTCATTTCTGCAAGTATGAGAAG 923
Oy      276 GYGIINIIIEuSeSeSeRIaAlaAspIIleAlaAlaThrValGlnHisThrMetAlaCys 295
Db      924 GGGCAAAATCCTGCTTTCAGACAGACACATTCCTGCACACAGTACAGCACAAATGGCATGT 983
Oy      296 HisLeuValVlyAspGlyThiHisArgAlaIleLeuPheCysGlyGlnArgAspLeuPro 315
Db      984 CATCTTGAAAAGAACACATCGGGGCTATTCTGTTTGTGAACAGAGAGACTTGTAACCT 1043
Oy      316 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAspNheTyrIleArgArg 335
Db      1044 CAAATTAATGCAGTACTGCTGTCATCTGCTGTGTGTCGCAAGTACTTCTATATCCGACA 1103
Oy      336 AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProPheArgLeu 355
Db      1104 GCTCGAAAATTTTAAACAAACGCAACACAGTCACTTGTTGTGTCCTCCCAACTA 1163
Oy      356 CysThrAspAspGlyIleMetIleAlaTrpAspGlyIleGlnArgLeuArgAlaGlyLeu 375
Db      1164 TGCACGTAAATGGCATTAATGATGTCATGCAATGGATGGATTAAGAACTAGCTGTGCTTG 1223
Oy      376 GlyIleLeuHisAspIleGlyGlyIleArgTyrGlnProLysCysProLeuGlyValAsp 395
Db      1224 GGCATTTTACATGACATAGAAAGCATCCGCTATGAACCAAAATGTCTCTTGGAGTAGAC 1283
Oy      396 IleSerLysGlyValGlyGlnAlaSerIleLysValProGlnLeuLysMetGlnIle 414
Db      1284 ATATCAAAAGAAAGTTGGAGAGACTTCATTAATAAGTAAACCAATTAAATAATGAGATA 1340

RESULT 9
US-10-012-140-6
; Sequence 6, Application US/10012140
; Publication No. US20030009017A1
; GENERAL INFORMATION:
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
; TITLE OF INVENTION: 14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
; TITLE OF INVENTION: THEROP
; FILE REFERENCE: 381552004800
; CURRENT APPLICATION NUMBER: US/10/012,140
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/246,768
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,772
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,185
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-140-6

Alignment Scores:
Pred. No.:      2.7e-243      Length:      1245
Score:          2088.00      Matches:      407
Percent Similarity: 99.03%      Conservative: 3
Best Local Similarity: 98.31%      Mismatches:  4
Query Match:    98.26%      Indels:      0
DB:             14          Gaps:        0

US-10-649-273-2 (1-414) x US-10-012-140-6 (1-1245)

Oy      1 MetLeuIleLeuThrLysThrAlaGlyValPhePheLysProSerLysArgLysValTyr 20
Db      1 ATGCTAAATCTTGACATAGACATGACAGAGTTTATTTTAAACATCAACAAAGAAAGTTTAT 60

```

QY 21 GluPheLeuArgSerPheAsnPheHisProGlyThrLeuPheLeuHisIleValLeu 40
 DB 61 GAATTTTAAAGAGTTTAAATTTTCACTTCGGAACACATATTTCTTCAATAAATGATATG 120
 QY 41 GlyTILEGluThrSerCysAspAspThrAlaAlaAlaValAspGluThrGlyAsnVal 60
 DB 121 GGAATTGAACTGATGATGATACAGCAGCTGCTGTGTGTGTAAGAACTGGAAATGTG 180
 QY 61 LeuGlyGluAlaIleHisSerGlnThrGluValHisIleuLeuSerThrGlyIleValPro 80
 DB 181 TTGGAGAAAGCAATACATTCCTCAACTGAAGTTCAATTTAAACAGGTGGATTTGTTCT 240
 QY 81 ProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGluAlaLeuSer 100
 DB 241 CCAGAGCTCAACAGCTTCAAGAGAAATATTCAGGAATAGTACAAAGAGCTTTTCT 300
 QY 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleLeuProGlyLeu 120
 DB 301 GCCAGTGAAGTCTCCCAAGTGAACCTCAGCAATTCACCACTAAACAGAGACTT 360
 QY 121 AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyIleLeuLeuLeu 140
 DB 361 GCTTTAAGCTGGAGAGGCTTATATCTTACAGCTGTAGAGCAGTAAAG 420
 QY 141 ProPheIleProIleHisIleMetGluAlaHisAlaLeuThrIleArgLeuThrAsnLeu 160
 DB 421 CCATTCATTCCTCATTCATATGAGAGCTCATGCACTTATAGTTACCAATAA 480
 QY 161 ValGluPheProPheLeuValLeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuVal 180
 DB 481 GTAGAAATTCCTTTTATGTTCTTTGATTTCTGAGAGTCACTGCTGTGGCATAGTT 540
 QY 181 GlnGlyValSerAspPheLeuLeuGlyIleSerLeuAspIleAlaProGlyAspMet 200
 DB 541 CAAGAGATTCAGATTTTCTGCTTCTTGAAAGCTTTTGACATAGCAACAGGTGACATG 600
 QY 201 LeuAspIleValAlaArgArgLeuSerLeuIleHisIleProGluCysSerThrMetSer 220
 DB 601 CTTGACAAAGGTGGCAAGACCTTCTTATTAACATCCAGAGTCTCCACATGAGT 660
 QY 221 GlyIleValAlaIleGluHisIleuAlaIleValGlnIleAsnArgPheHisPheAspIleLeu 240
 DB 661 GTGGGAAAGCCATAGAACATTTGGCCAAACAGAAATATATTCATTTGACATCAA 720
 QY 241 ProProLeuHisHisIleValAsnCysAspPheSerPheThrGlyLeuGlnHisValThr 260
 DB 721 CTTCCCTTGATCATGCTAAATAATTTGATTTTCTTACTGACATTCACACAGTTACT 780
 QY 261 AspIleValIleMetIleValGlyGluGluGluGluGluGluGluGluGluGluGluGlu 280
 DB 781 GATTAATAATATATGAAACAGGAAACAGAGAGATTTGAGAGGGGCAATCTGTCT 840
 QY 281 SerAlaIleAspIleAlaIleAlaThrValGlnHisIleThrMetAlaCysHisIleValIleArg 300
 DB 841 TCGACAGAGCATTTGCTGCCACAGTACAGCAACAAATGCAATGTCATCTTGTAAGA 900
 QY 301 ThrHisArgAlaIleLeuPheCysLeuGlnArgAspLeuLeuProGlnAsnAsnAlaVal 320
 DB 901 ACACATCGGGCTATTTCTTTTGTAGCAGAGAGCTTTTACCTCAAAATATATGCAATA 960
 QY 321 LeuValAlaSerGlyGlyValAlaSerAsnPheThrIleArgArgAlaLeuGluIleLeu 340
 DB 961 CTGGTTCATCTGGTGTGTGCGCAAGTAACTTATATCCGACAGAGCTCTGGAATTTTA 1020
 QY 341 ThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLeuCysThrAspAsnIle 360
 DB 1021 ACAAACCAACACAGTGCATTTGTGTCTCTCCACAGATATGACATGATATATGAC 1080
 QY 361 IleMetIleAlaTPAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAsn 380
 DB 1081 ATTATGATTCATGAGATGATGATTAAGAACTACCTGCTGGCTTGGCAATTTTACATAC 1140
 QY 381 IleGluGlyIleArgTyrGluProLeuCysAspProLeuGlyAlaAspIleSerIleGluVal 400

DB 1141 ATAAAGGCATCCCTATGAAACAAATGCTCTTGGAGTACATATACAAAGAGTT 1200
 QY 401 GlyGluAlaSerIleIleValProGlnLeuIleuMetGluIle 414
 DB 1201 GGAAGCTTCCATTAAGTACCAATTAATAATGAGATA 1242
 RESULT 10
 US-10-012-140-4
 ; Sequence 4, Application US/10012140
 ; Publication No. US2003009017A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Leidy, Kevin R.
 ; APPLICANT: Kapeller-Libermann, Rosana
 ; APPLICANT: Glucksmann, Maria A.
 ; TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
 ; TITLE OF INVENTION: 14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: 38155204900
 ; CURRENT APPLICATION NUMBER: US/10/012,140
 ; PRIOR FILING DATE: 2001-11-08
 ; PRIOR APPLICATION NUMBER: 60/246,768
 ; PRIOR FILING DATE: 2000-11-08
 ; PRIOR APPLICATION NUMBER: 60/246,772
 ; PRIOR FILING DATE: 2000-11-08
 ; PRIOR APPLICATION NUMBER: 60/249,185
 ; PRIOR FILING DATE: 2000-11-15
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 1820
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (146)...(1390)
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(1820)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-012-140-4
 Alignment Scores:
 Pred. No.: 4,91e-243 Length: 1820
 Score: 2088.00 Matches: 407
 Percent Similarity: 99.03% Conservative: 3
 Best Local Similarity: 98.31% Mismatches: 4
 Query Match: 98.26% Indels: 0
 DB: 14 Gaps: 0
 US-10-649-273-2 (1-414) x US-10-012-140-4 (1-1820)
 QY 1 MetLeuIleLeuThrIleThrAlaGlyValPhePheIleProSerIleArgIleValTyr 20
 DB 146 ATGCTAATCTTGACATCAAGCTGACAGAGTTTAAAAACATCAAAAAAGAAAGTTAT 205
 QY 21 GluPheLeuArgSerPheAsnPheHisProGlyThrLeuPheLeuHisIleValLeu 40
 DB 206 GAATTTTAAAGAGTTTAAATTTTCACTTCGGAACACATATTTCTTCAATAAATGATATG 265
 QY 41 GlyTILEGluThrSerCysAspAspThrAlaAlaAlaValAspGluThrGlyAsnVal 60
 DB 266 GGAATTGAACTGATGATGATACAGCAGCTGCTGTGTGTGTAAGAACTGGAAATGTG 325
 QY 61 LeuGlyGluAlaIleHisSerGlnThrGluValHisIleuLeuSerThrGlyIleValPro 80
 DB 326 TTGGAGAAAGCAATACATTCCTCAACTGAAGTTCAATTTAAACAGGTGGATTTTCT 385
 QY 81 ProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGluAlaLeuSer 100
 DB 386 CCAGAGCTCAACAGCTTCAAGAGAAATATTCAGGAATAGTACAAAGAGCTTTTCT 445
 QY 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleLeuProGlyLeu 120

Db 446 GCCAGTGGAGTCTCTCCAGTGAACCTCGACCAATGCACTACCAATTAACCAAGGACTT 505
Qy 121 AAlaSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuLys 140
Db 506 GCTTTAAGCTGGGAGGAGGCTTATCATTTAGCTTACAGCTGGAGGAGGCTTAAAAAG 565
Qy 141 ProPheIleProIleHisIleMetGlnAlaHisAlaLeuThrIleArgLeuThrAsnLys 160
Db 566 CCATTCAATCCCATTCATCATATGAGAGGCTCATCTACTATTAGGTGACCAATTA 625
Qy 161 ValGluPheProPheLeuValIleLeuIleSerGlyGlyHisIleCysLeuLeuAlaLeuVal 180
Db 626 GTAGAATTCCTTTTAAAGTTCTTTTAAATTCGAGAGTCACTGCTGGGCAATTAGTT 685
Qy 181 GlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMet 200
Db 686 CAAAGAGTTCAAGTTTCTGCTCTTGGAAAGCTTTGGACATAGCACCGAGTGACATG 745
Qy 201 LeuAspLysValAlaArgArgLeuSerLeuIleLysIleProGluCysSerThrMetSer 220
Db 746 CTGGACAAAGTGGCAAGAGCTTTCTTTAATAAATCATCCAGAGTGCTCCACCATGAGT 805
Qy 221 GlyGlyLysAlaIleGlyHisLeuAlaLysGlnGlyAsnArgPheHisPheAspIleLys 240
Db 806 GGTGGGAAAGCCATAGAACATTTGGCCAAACAAAGAAATAGATTTCAATTTGACATCAA 865
Qy 241 ProProLeuHisIleAlaLysAsnCysAspPheSerPheThrGlyLeuGlnHisValThr 260
Db 866 CTTCCCTTGATCATAGCTAAATAATTTGATTTTCTTTTACTGGACCTTCAACGTTACT 925
Qy 261 AspLysIleIleMetLysGlyLysGluGluGlyIleGluLysGlyGlnIleLeuSer 280
Db 926 GATTAATAATATGATAAACAAGAGAAACAAGAGATTTGAGAGGGGCAAAATCTGCT 985
Qy 281 SerAlaAlaAspIleAlaAlaThrValGlnHisIleThrMetAlaCysHisLeuValLysArg 300
Db 986 TCAGCAGCAGACATTTGCTGCACAGTACAGACACAAATGGCATGTCTTTGTGAAAAA 1045
Qy 301 ThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuProGlnAsnAsnAlaVal 320
Db 1046 AACAATCGGGCTATCTGTTTGTGAGCAGAGACATTTGTTACTCAAAATAAGCGAGTA 1105
Qy 321 LeuValAlaSerGlyValAlaSerAsnPheThrIleArgArgAlaLeuGlnIleLeu 340
Db 1106 CTGGTTGCATCTGCTGTGTGTCAGAGTAACTTCTATATCCGACAGCTCTGGAATTTTA 1165
Qy 341 ThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeuCysThrAspAsnGly 360
Db 1166 ACAAAACGCAACACAGTGCACCTTTGTTGTGCTCCACAGACTATGCACATGATATGGC 1225
Qy 361 IleMetIleAlaThrPangGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAsp 380
Db 1226 ATTATATATTCATGAGATGATGATTAAGAACTAGTCTGGCTGGCATTTTTCATGAC 1285
Qy 381 IleGlnGlyIleArgGlyArgGluProLysCysProLeuGlyValAspIleSerLysGlyVal 400
Db 1286 ATTAGAAGGCTCCGCTATGAAACCAAAATGCTCTTTGAGTAGCAATATCAAAAGAGTT 1345
Qy 401 GlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 414
Db 1346 GGAAGAGCTTCCATTAAGAACCAATTAATAATGAGAGTA 1387

RESULT 11
US-10-094-749-400
; Sequence 400, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI

APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUKIO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: YAMASHITA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOMYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 400
LENGTH: 2208
TYPE: DNA
ORGANISM: Homo sapiens
US-10-094-749-400

Alignment Scores:
Pred. No.: 2,266-225 Length: 2208
Score: 1944.00 Matches: 386
Percent Similarity: 93.24% Conservative: 0
Best Local Similarity: 93.24% Mismatches: 4
Query Match: 91.48% Indels: 24
Gaps: 17

US-10-649-273-2 (1-414) x US-10-094-749-400 (1-2208)

Qy 1 MetLeuIleLeuThrIleThrAlaGlyValPhePheLysProSerLysArgLysValTyr 20
Db 344 ATGCAATCTTGACTAGCACTGACAGAGCTTTTAAACATCAAAAGGAAAGTTAT 403
Qy 21 GluPheLeuArgSerPheAsnPheHisProGlyThrLeuPheLeuHisLysIleValLeu 40
Db 404 GAATTTTAAAGTTTAAATTTTATTTTCTGGAACATTTCTTCAATAATTAAGATTG 463
Qy 41 GlyIleGluThrSerCysAspAspThrAlaAlaAlaValAspGluThrGlyAsnVal 60
Db 464 GGAATTAAGAACTAGTTGATGATACAGACAGCTGCTGTGTGATGAACCTGGAATGTG 523
Qy 61 LeuGlyGluAlaIleHisIleSerGlnThrGluValHisLeuLysThrGlyGlyIleValPro 80
Db 524 TTGGAGAGAGAAATACATTTCCCAACTGAAGTTCTTTAAAAACAGTGGAGTTGCT 583
Qy 81 ProAlaIleGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGlnAlaLeuSer 100
Db 584 CAGCAGCTTCAACACCTTCAACAGAGAAATATTCAACAAATATGTAACAGAGCTCTTTT 643
Qy 101 AlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleLysProGlyLeu 120
Db 644 GCCAGTGAAGTCTGCCAAGTGAAGCTCTCAGAAATTTGCACTACATTAACCAAGACTT 703
Qy 121 AlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuLysLys 140
Db 704 GCTTTAAGCTGGGAGGCTTATCATTTAGCTTACAGCTGGAGGAGGCTTAAAAAG 763
Qy 141 ProPheIleProIleHisIleMetGlnAlaHisAlaLeuThrIleArgLeuThrAsnLys 160
Db 764 CCATTCAATCCCATTCATCATATGAGAGGCTCATCTACTATTAGGTGACCAATTA 823
Qy 161 ValGluPheProPheLeuValIleLeuIleSerGlyGlyHisIleCysLeuLeuAlaLeuVal 180

```

Db      824 GTAGAAATTCCTTTTATGTTCTTTGATTTCTGGAGCTCATCTCTGTTGGCAATTAGTT 883
Qy      181 GInG1yValSerAspPheLeuLeuLeuG1yLysSerLeuAspIleAlaProG1yAspMet 200
Db      884 CAAAGAGTTTCAGATTTTCTGCTTCTTGAAAGCTTTTGACATACAGACACAGGAGCATG 943
Qy      201 LeuAspLysValAlaArgArgLeuSerLeuIleLysIleProG1yCysSerThimetSer 220
Db      944 CTTGACAAAGGTGCAGAAAGACCTTCTTATATAAATCATCCAGAGTCTCCACCATGAGT 1003
Qy      221 G1yG1yLysAlaIleG1uHisLeuAlaLysG1nG1yLysArgPheHisPheAspIleLys 240
Db      1004 GGTGGAAAGCCCTAGAAACATTTGGCCAAACAAAGAAATATATTTCAATTTGACATCAAA 1063
Qy      241 ProProLeuHisHisAlaLysAsnCysAspPheSerPheThrg1yLeuG1nHisValThr 260
Db      1064 CTTCCCTTGATCATGCTATTAATAATTTGATTTTCTTTACTTGACCTTCAACAGCTTACT 1123
Qy      261 AspLysIleIleMetLysLysG1uLysG1uG1uG1yIleG1uLysG1yG1nIleLysSer 280
Db      1124 GATTAATATATATATGAAAAGAAAAGAGAAAGATTTGAAAGGGGCAAAATCTGTCT 1183
Qy      281 SerAlaAlaAspIleAlaAlaThrValG1nHisThrMetAlaCysHisLeuValLysArg 300
Db      1184 TCAGCAGACAGACATTTGCTGCACAGTACAGACACAAATGCGATGTCACTTTGAAAAGA 1243
Qy      301 ThrHisArgAlaIleLeuPheCysLysG1yArgArgAspLeuProG1nAspAsnAlaVal 320
Db      1244 ACAATCCGGGCTATTTCTGTTTGTAAACAGAGACATTTGTAATCTCAAAATTAATGAGTA 1303
Qy      321 LeuValAlaSerG1yG1yValAlaSerAspPheThrg1yLeuArgArgAlaLeuG1uIleLeu 340
Db      1304 CTGCTTGATCTGTGGTGTGTGCGCAAGTAACTTCTGTATCCGCAAGCTCTGGAATTTTA 1363
Qy      341 ThrAsnAlaThrG1nCysThrLeuLeuCysProProArgLeuG1yCysThrAspAsnG1y 360
Db      1364 ACAAGCGCAACAGTGCACATTTGTGTCTCTCCACAGATATGACATGATATATGCG 1423
Qy      361 IleMetIleAlaTrpAsnG1yIleG1uArgLeuArgAlaG1yLeuG1yIleLeuHisAsp 380
Db      1424 ATATGATTCGA-----TGATGCTCTTGAGTAAATATCAATCAAAAGAGTT 1435
Qy      381 IleG1uG1yIleArgTrpG1uProLysCysProLeuG1yValAspIleSerLysG1uVal 400
Db      1436 -----TGATGCTCTTGAGTAAATATCAATTAATAATGAGATA 1471
Qy      401 G1yG1uAlaSerIleLysValProG1nLeuLysMetG1uIle 414
Db      1472 GGAGAGCTTCCATTAATAAGTACCAATTAATAATAATGAGATA 1513

RESULT 12
US-10-723-860-7447
; Sequence 7447, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natascha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NFUS01
; CURRENT APPLICATION NUMBER: US/10/723, 860
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7447
; LENGTH: 2890
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: (646)..(657)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-7447

Alignment Scores:
Pred. No.: 3,45e-225 Length: 2890
Score: 1944.00 Matches: 386
Percent Similarity: 93.24% Conservatave: 0
Best Local Similarity: 93.24% Mismatches: 4
Query Match: 91.48% Indels: 24
DB: 20 Gaps: 1

US-10-649-273-2 (1-414) x US-10-723-860-7447 (1-2890)
Qy      1 MetLeuIleLeuThrLysThrAlaG1yValPhePheLysProSerLysArgLysValThr 20
Db      1001 ATGCTAATCTTGACTTAAGACTGACAGACTTTTATTAACATCAAAAAGAAATTTAT 1060
Qy      21 GluPheLeuArgSerPheAsnPheHisProG1yThrLeuPheLeuHisLysIleValLeu 40
Db      1061 GAATTTTAAGAAGTTTAAATTTTCACTCGAAACATATTCTTCATTAATAATAGTATG 1120
Qy      41 G1yIleG1uThrSerCysAspAspThrAlaAlaValAlaAspG1uThrG1yLysVal 60
Db      1121 GGAATTTGAATAGTGTGATGATACAGACACTGCTGTGGTGAATGAACAGGAATGTG 1180
Qy      61 LeuG1yG1uAlaIleHisSerG1nThrG1yValHisLeuLysThrG1yG1yIleValPro 80
Db      1181 TTGGAGAAAGCAATACATTTCCCAACTGAAGTTCACTTTAAACAGGGGGAATTTCTCT 1240
Qy      81 ProAlaAlaG1nG1nLeuHisArgG1uAsnIleG1nArgIleValG1nG1uAlaLeuSer 100
Db      1241 CCAGAGCTCAACAGCTTCAACAGAAATTTTCAACGAATGTCAAGAAAGCTCTTCT 1300
Qy      101 AlaSerG1yValSerProSerAspLeuSerAlaIleAlaThrIleLysProG1yLeu 120
Db      1301 GCCAGTGAAGTCTCTCCAAAGTGAACCTTCAGAAATTTGAACATCAATTAACCAAGACTT 1360
Qy      121 AlaLeuSerLeuG1yValG1yLeuSerPheSerLeuG1nLeuValG1yG1nLeuLysLys 140
Db      1361 GCTTTAAGCTTGGAGTGGCTTATCATTTAGCTTACAGCTGTGAGCAAGTTAAAG 1420
Qy      141 ProPheIleProIleHisIleMetG1uAlaHisAlaLeuThrIleArgLeuThrAsnLys 160
Db      1421 CCATTCATTTCCCATTCATCATATGAGGCTCATGACATTAATAGTTGACCAATAAA 1480
Qy      161 ValGluPheProPheLeuValLeuLeuIleSerG1yG1nHisCysLeuLeuAlaLeuVal 180
Db      1481 GTAGAAATTCCTTTTATGTTCTTTGATTTCTGAGAGTCACTGTCTGTGGCATTAGTT 1540
Qy      181 GInG1yValSerAspPheLeuLeuLeuG1yLysSerLeuAspIleAlaProG1yAspMet 200
Db      1541 CAAAGAGTTTCAAGATTTTCTGCTTCTTGAAGTCTTTGAGATGACACAGGTACATG 1600
Qy      201 LeuAspLysValAlaArgArgLeuSerLeuIleLysIleProG1yCysSerThimetSer 220
Db      1601 CTGCAAGGTGCGCAAGAACTTCTTTATATAAATCCAGAGTCTCCACCATGACT 1660
Qy      221 G1yG1yLysAlaIleG1uHisLeuAlaLysG1nG1yLysArgPheHisPheAspIleLys 240
Db      1661 GGTGGAAAGCCATAGAACATTTTGCCCAACAGAAATATAGTTCAATTTTCAATCAAA 1720
Qy      241 ProProLeuHisHisAlaLysAsnCysAspPheSerPheThrg1yLeuG1nHisValThr 260
Db      1721 CTTCCCTTGATCATGCTAATAAATTTGATTTTCTTTTACATGCACTTCAACACCTTACT 1780
Qy      261 AspLysIleIleMetLysLysG1uLysG1uG1uG1yIleG1uLysG1yG1nIleLysSer 280
Db      1781 GATTAATAATATATATAAAGAAAAGAAAGAAAGATTTGAGAAAGGAAATCCCTGTCT 1840
Qy      281 SerAlaAlaAspIleAlaAlaThrValG1nHisThrMetAlaCysHisLeuValLysArg 300
Db      1841 TCAGCAGCAGACATTTGCTGCACAGTACAGACACAAATGCGATGTCACTTTGAAAAGA 1900

```

QY 301 ThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuProGlnAsnAlaVal 320
DB 1901 ACACATGGGCTATTCTGTTTGTAGCAGAGACTGTGTAACCAATTAATGCACTA 1960
QY 321 LeuValAlaSerGlyValAlaSerAspPheTyrIleArgAlaIleGluIleLeu 340
DB 1961 CTGTGGTGCATCTGGTGTGTGCGAAGTAATCTGTATCCGCAAGCTCTGGAATTTTA 2020
QY 341 ThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeuCysThrAspAsnGly 360
DB 2021 ACMAAGCCACACAGTGCATCTTGTGTGTCTCTCCAGACTATGCACTGATATGCGC 2080
QY 361 IleMetIleAlaTyrAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAsp 380
DB 2081 ATTATGATTCGA----- 2092
QY 381 IleGluGlyIleArgTyrGluProLysCysProLeuGlyValAlaPylIleSerLysGluVal 400
DB 2093 -----TGATGTCTCTCTGGAGTAGACATATCAAAAGAGTT 2128
QY 401 GlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 414
DB 2129 GGAGAGCTTCATMAAAGTACCAATMAAATGAGATA 2170
RESULT 13
US-10-120-988-177
; Sequence 177, Application US/10120988
; Publication No. US20030219745A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryje
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyun
; APPLICANT: Wang, Dunrui
; APPLICANT: Dzmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: PL_FL_genes Version 2.0
; SEQ ID NO 177
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (205)..(1305)
US-10-120-988-177
Alignment Scores:
Pred. No.: 1 07e-201 Length: 1416
Score: 1747.00 Matches: 340
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.71% Mismatches: 0
Query Match: 82.21% Indels: 0
DB: 17 Gaps: 0
US-10-649-273-2 (1-414) x US-10-120-988-177 (1-1416)
QY 74 LysThrGlyGlyIleValProProAlaAlaGlnLeuHisArgGluAsnIleGlnArg 93
DB 280 AGAACAAGGTGGATTGTTCTCCGACGACTCAACAGCTTCCACAGAGAAATATTCAACGA 339
QY 94 IleValGlnGluAlaLeuSerAlaSerGlyValSerProSerAspLeuSerAlaIleAla 113
DB 340 ATAGTACAAAGAGCTTCTTCTGCGACAGTGTCTCTCAATGACCTCTCAGCAATTCGA 399
QY 114 ThrThrIleLysProGlyLeuAlaLeuSerIleuGlyValGlyLeuSerPheSerIleuGln 133

DB 400 ACTACCAATMAAAGAGCTTGTAGCCCTGGAGAGGCTTATCATTTAGCTTACG 459
QY 134 LeuValGlyGlnLeuLysLysProPheIleProIleHisMetGluAlaHisAlaLeu 153
DB 460 CTGTGGAGACGATTMAAAGCCATTCTTCCATTATCATATGAGGCTCATGCATCT 519
QY 154 ThrIleArgLeuThrAsnLysValGluPheProPheLeuValLeuLeuIleSerGlyGly 173
DB 520 ACTATTAGTGGACCAATTAAGTAGAATTTCCCTTTTATGTTCTTTGATTTCTGGAGGT 579
QY 174 HisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeu 193
DB 580 CACTGTCTGTGGCATTTGTTCAAGAGTTTCAGATTTTCCTTCTTGGAAAGCTTTTG 639
QY 194 AspIleAlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerIleLysHis 213
DB 640 GACATGACCAAGGTGATGCTTTCACAAAGGTGGCAAGAGACTTTCTTATTAATAACAT 699
QY 214 ProGluCysSerThrMetSerGlyGlyLysAlaIleGlnHisLeuAlaLysGlnGlyAsn 233
DB 700 CCAAGTCTCCACATGAGTGTGGGAAAGCCATGAAACATTTGGCCAAACAGGAANT 759
QY 234 ArgPheHisPheAspIleLysProProLeuHisHisAlaLysAsnCysAspPheSerPhe 253
DB 760 AGATTTCATTTGACATCAAACTCCCTTGATCATGCTAAAAATTTGATTTTCTTTT 819
QY 254 ThrGlyLeuGlnHisValThrAspLysIleIleMetLysGlnLysGlnGluGlyIle 273
DB 820 ACTGACTTCAACACAGTTACTGATTAATTAATGAAAGAAAGAAAGAGAGAGTATT 879
QY 274 GluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaIleThrValGlnHisThrMet 293
DB 880 GAGAAAGGCAATCTCTGTTCTTCCAGCAGACATTCCTGCCACGTACAGACCAAG 939
QY 294 AlaCysHisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeu 313
DB 940 GCATGTCACTTGTGAAAGAAACATCGGGCTATTCTGTTTGTAGCAGAGACTTG 999
QY 314 LeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaLysSerAspPheTyrIle 333
DB 1000 TTACTCTAAATTAAGCAGTCTGTGTTCATCTGTGTGTGTGCGCAAGTAATCTTATATC 1059
QY 334 ArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProPro 353
DB 1060 CCGAGAGCTCTGGAATTTTAACCAACGACACAGTGCATTTGTGTCTCTCC 1119
QY 354 ArgLeuCysThrAspAsnGlyIleMetIleAlaTyrAsnGlyIleGluArgLeuArgAla 373
DB 1120 AGACTATGCATGATTAATGCGATTATGATGATGATGATGATGATGATGATGATGATGAT 1179
QY 374 GlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGluProLysCysProLeuGly 393
DB 1180 GGCCTGGCATTTTACATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 1239
QY 394 ValAspIleSerLysGluValGlyGluAlaSerIleLysValProGlnLeuLysMetGlu 413
DB 1240 GTAGCATATCAAAAGAGGTGGAGAGCTTCCATTAAGTACCAATTAATAATGAG 1299
QY 414 Ile 414
DB 1300 ATA 1302
RESULT 14
US-10-067-443-23
; Sequence 23, Application US/10067443
; Publication No. US20030082782A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE HIGHLY EXPRESSED IN
; FILE REFERENCE: D0073 NP
; CURRENT APPLICATION NUMBER: US/10/067,443

```
/ CURRENT FILING DATE: 2002-02-05
/ PRIOR APPLICATION NUMBER: US 60/266,518
/ PRIOR FILING DATE: 2001-02-05
/ PRIOR APPLICATION NUMBER: US 60/282,814
/ PRIOR FILING DATE: 2001-04-10
/ NUMBER OF SEQ ID NOS: 71
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 23
/ LENGTH: 1526
/ TYPE: DNA
/ ORGANISM: homo sapiens
US-10-067-443-23

Alignment Scores:
Pred. No.: 1,41e-157 Length: 1526
Score: 1385.00 Matches: 267
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65,18% Indels: 0
DB: 14 Gaps: 0

US-10-649-273-2 (1-414) x US-10-067-443-23 (1-1526)
QY 148 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 167
DB 1 ATGGAGGCTCATGCACTTACTATTAGTTGACCAATAAGTAGAATTTCTTTTAGTT 60
QY 168 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 187
DB 61 CTTTGAATTTCTGGAGGCTCACTGCTGTGGCATTAGTTCAAGAGATTTTCGATTTCTG 120
QY 188 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 207
DB 121 CTTCTGGAAGGCTTTTGACATAGCACAGGTACAGCTTGACAGAGTGGAAGA 180
QY 208 LeuSerLeuIleLysHisProGlyCysSerThrMetSerGlyGlyLysAlaIleGlnHis 227
DB 181 CTTTCTTAATAAACAATCCAGAGTCTCCACCATGAGTGGGAAGCCATGAAACAT 240
QY 228 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProPheLeuHisAlaLys 247
DB 241 TTGGCCAAACAAGGAATATGATTTCAATTTGACATCAACCTCCCTTCATCATGTA 300
QY 248 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 267
DB 301 AATTGATTTTCTTTTACTGACCTTCAACACCTTCTGATTAATAATATGAAAAAG 360
QY 268 GluLysGlnGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaAla 287
DB 361 GAAAAAGAGAGGATATGAGAGGGGCAATCTGTCTTCAGCAGACATTTGCTGCC 420
QY 288 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 307
DB 421 ACAGTACAGACACAAATGGCATGTCATCTTGTGAAAAGAACACATCGGCTATTTGTT 480
QY 308 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyVal 327
DB 481 TGTAAAGAGAGAGCTTGTAACTCAAAATATATGACAGTACTGCTGCTGCTGCTG 540
QY 328 AlaSerAsnPheThrIleArgArgAlaLeuGlnIleLeuThrAsnAlaIleThrGlyCysThr 347
DB 541 GCAAGTAACCTTATATCCGACAGCTCTGGAATTTTAAACAAACGCAACACAGTCACT 600
QY 348 LeuLeuCysProPheProPheArgLeuCysThrAspAsnGlyIleMetIleAlaTriPheAsnGly 367
DB 601 TTGTGTGCTCTCTCCACAGCTATGCACTGATTAATGCAATTAATGATTCATGAAATGT 660
QY 368 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGlnGlyIleArgTyrGln 387
DB 661 ATTGAAGAAGACTACGTGCTGCTGGGCAATTTATCATGACATGAAGGATCGGCTATGA 720
QY 388 ProLysCysProLeuGlyValAspIleSerLysGlnValGlyGluAlaSerIleLysVal 407
```

```
DB 721 CCAAAATGCTCTTGGAGTACATATCAAAAGAAGTTGGAGAGCTTCCATTAAGTA 780
QY 408 ProGlnLeuLysMetGluIle 414
DB 781 CCACATTTAAAAATGAGAGATA 801

RESULT 15
US-10-649-273-23
/ Sequence 23, Application US/10649273
/ Publication No. US20040043407A1
/ GENERAL INFORMATION:
/ APPLICANT: Bristol-Myers Squibb Company
/ TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE, MP-1
/ FILE REFERENCE: D0073 CNT
/ CURRENT APPLICATION NUMBER: US/10/649,273
/ CURRENT FILING DATE: 2003-08-27
/ PRIOR APPLICATION NUMBER: US 60/266,518
/ PRIOR FILING DATE: 2001-02-05
/ PRIOR APPLICATION NUMBER: US 10/067,443
/ PRIOR FILING DATE: 2002-02-05
/ PRIOR APPLICATION NUMBER: US 60/282,814
/ PRIOR FILING DATE: 2001-04-10
/ NUMBER OF SEQ ID NOS: 71
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 23
/ LENGTH: 1526
/ TYPE: DNA
/ ORGANISM: homo sapiens
US-10-649-273-23

Alignment Scores:
Pred. No.: 1,41e-157 Length: 1526
Score: 1385.00 Matches: 267
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65,18% Indels: 0
DB: 18 Gaps: 0

US-10-649-273-2 (1-414) x US-10-649-273-23 (1-1526)
QY 148 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 167
DB 1 ATGGAGGCTCATGCACTTACTATTAGTTGACCAATAAGTAGAATTTCTTTTAGTT 60
QY 168 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 187
DB 61 CTTTGAATTTCTGGAGGCTCACTGCTGTGGCATTAGTTCAAGAGATTTTCGATTTCTG 120
QY 188 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 207
DB 121 CTTCTGGAAGGCTTTTGACATAGCACAGGTACAGCTTGACAGAGTGGAAGA 180
QY 208 LeuSerLeuIleLysHisProGlyCysSerThrMetSerGlyGlyLysAlaIleGlnHis 227
DB 181 CTTTCTTAATAAACAATCCAGAGTCTCCACCATGAGTGGGAAGCCATGAAACAT 240
QY 228 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProPheLeuHisAlaLys 247
DB 241 TTGGCCAAACAAGGAATATGATTTCAATTTGACATCAACCTCCCTTCATCATGTA 300
QY 248 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 267
DB 301 AATTGATTTTCTTTTACTGACCTTCAACACCTTCTGATTAATAATATGAAAAAG 360
QY 268 GluLysGlnGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaAla 287
DB 361 GAAAAAGAGAGGATATGAGAGGGGCAATCTGTCTTCAGCAGACATTTGCTGCC 420
QY 368 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 387
DB 421 ACAGTACAGACACAAATGGCATGTCATCTTGTGAAAAGAACACATCGGCTATTTGTT 480
QY 388 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyVal 327
```


D _b	481	TGTAAGCAGAGACTTGTACTCTCAAAATAATGCAGTACTGGTGCATCTGGTGTTC	540
Q _y	328	AlaSerAsnPhenylTlLeuArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThr	347
D _b	541	GCAAGTAATCTCTAATATCCGACAGACTCTGGAAATTTTAAACAAACGCAACACAGTGCACCT	600
Q _y	348	LeuLeuCysProProProArgLeuCysThrAspAsnGlyTlMetIleAlaTPrAsnGly	367
D _b	601	TGTGTGTCTCTCCCTCCAGACTATGCCTGATAATGGCATTTATGATTCACATGAAATGGT	660
Q _y	368	IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyTlLeuArgTyrGlu	387
D _b	661	ATTGAAACACTACGTCGTGGCTTTTACATTTTACATGACATGAAAGCATTCGCTATGAA	720
Q _y	388	ProLysCysProLeuGlyValAspIleSerTyrGluValGlyGluAlaSerIleLysVal	407
D _b	721	CCAAATGTCTCTCTTGAGATGACATATCAAAAAGAAAGTGGAGAGACTTCATATAAAGTA	780
Q _y	408	ProGlnLeuLysMetGluIle	414
D _b	781	CCACATTTAAATAATGAGATTA	801

Search completed: June 17, 2005, 08:03:08
Job time : 2842.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: June 16, 2005, 15:19:45 ; Search time 3193.26 Seconds
(Without alignments)
4051.513 Million cell updates/sec

Title: US-10-649-273-2_COPY_148_414
Perfect score: 1385
Sequence: 1 MEAHATIRLTKVKEPFLV.....DISKEVGEASIKVPQLMEI 267

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q/cgna2_1/USPTO.spool_p/US10649273/runat.15062005.111416.6030/app_query.fasta_1.1429
-DB=GenEmbl -OFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LCOPT=0 -LCOPEXT=0
-INITs=bits -START=1 -END=-1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=us10649273 @CGN 1.1 7742 @runat.15062005.111416.6030 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb ba:.*
2: gb htg:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1385	100.0	1416	AR541929	AR541929 Sequence
2	1385	100.0	1526	AR428809	AR428809 Sequence
3	1385	100.0	1908	BC011904	BC011904 Homo sapi
4	1385	100.0	2197	AR428803	AR428803 Sequence

5	1358.5	98.1	1387	6	AR428808	AR428808 Sequence
6	1358.5	98.1	1387	9	HS295148	AU795148 Homo sapi
7	1348	97.3	1245	6	AX664697	AX664697 Sequence
8	1348	97.3	1820	6	AX664695	AX664695 Sequence
9	1231	88.9	1844	10	BC058172	BC058172 Mus muscu
10	1224	88.4	1017	10	BC038910	BC038910 Mus muscu
11	1204	86.9	2208	6	AX713716	AX713716 Sequence
12	1204	86.9	2208	9	AK055441	AK055441 Homo sapi
13	1116	80.6	1546	10	BC078974	BC078974 Rattus no
14	950.5	68.6	1109	10	BC051211	BC051211 Mus muscu
15	939	67.8	1522	5	BX934991	BX934991 Gallus ga
16	938	67.7	860	5	BX930694	BX930694 Gallus ga
17	934	67.4	1558	5	BX930963	BX930963 Gallus ga
18	831.5	60.0	84115	9	AC013468	AC013468 Homo sapi
19	824.5	59.5	14364	6	AR428807	AR428807 Sequence
20	666	48.1	249601	2	AC114153	AC114153 Rattus no
21	666	48.1	308652	2	AC121478	AC121478 Rattus no
22	652.5	47.1	256751	10	AC122925	AC122925 Mus muscu
23	610	44.0	875	6	CQ721898	CQ721898 Sequence
24	488.5	35.3	1474	3	AK113378	AK113378 Ciona int
25	477.5	34.5	121251	5	AL591593	AL591593 Zebrafish
26	435	31.4	117322	5	AL672217	AL672217 Zebrafish
27	415.5	30.0	1576	3	AY051882	AY051882 Drosophi
28	415.5	30.0	1601	6	CQ606432	CQ606432 Sequence
29	415.5	30.0	3656	6	CQ606431	CQ606431 Sequence
30	415.5	30.0	14679	2	AC018262	AC018262 Drosophi
31	415.5	30.0	180263	3	AC010671	AC010671 Drosophi
32	415.5	30.0	207432	3	AE003513	AE003513 Drosophi
33	408	29.5	1672	8	AK070912	AK070912 Oryza sat
34	407.5	29.4	1443	8	AY024338	AY024338 Arabidops
35	407.5	29.4	1474	8	AY117283	AY117283 Arabidops
36	407.5	29.4	1567	8	AY063864	AY063864 Arabidops
37	401.5	28.0	1557	8	AY084577	AY084577 Arabidops
38	389	26.1	110000	2	AF006501_08	Continuation (9 of
39	369	26.6	260271	1	AE017258	AE017258 Wolbachia
40	356.5	25.7	33800	1	SMES91792	AL591792 Sinorhizo
41	356	25.7	20021	10	AF367967S1	AF367967 Mus muscu
42	356	25.7	179252	10	AF131205	AF131205 Mus muscu
43	352	25.4	349116	1	AB003003	AP003003 Mesorhizo
44	349	25.2	110000	1	AE017197_01	Continuation (2 of
45	342	24.7	756	6	AX685201	AX685201 Sequence

ALIGNMENTS

RESULT 1
LOCUS AR541929 1416 bp DNA
DEFINITION Sequence 177 from patent US 6743619.
ACCESSION AR541929
VERSION AR541929.1 GI:53934009
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1416)
Tang,Y.T., Zhou,P., Goodrich,R., Liu,C., Asundi,V., Ren,F.,
AUTHORS Zhang,J., Zhao,Q.A., Yang,Y., Xue,A.J., Wehrman,T., Wang,J.-R.,
Wang,D. and Drmanac,R.T.
TITLE Nucleic acids and polypeptides
JOURNAL Patent: US 6743619-A 177.01-JUN-2004;
FEATURES
source Location/Qualifiers
1..1416
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.37e-126
Score: 1385.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Length: 1416
Matches: 267
Mismatches: 0
Indels: 0

RESULT 3
BC011904
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCES
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT

BC011904 1908 bp mRNA linear PRI 23-DEC-2003
Homo sapiens O-sialoglycoprotein endopeptidase-like 1, mRNA (CDNA
clone MGC:20293 IMAGE:4121450), complete cds.
BC011904
BC011904.2 GI:40225818
MGC.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1908)
Straussberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.T., Wang, J., Heide, F.,
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Tothyluki, S.,
Carinci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwen, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Sanchez, J., Helton, E., Ketelem, M., Madan, A., Young, A.C., Rodriguez, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,
Gschneer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1908)
Straussberg, R.
Direct Submission
Submitted (30-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Dec 19, 2003 this sequence version replaced gi:15080281.
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, J., Gupta, J., Hachiguchi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stancir, S., Thomas, P.J., Touchman, J.W.,
Tsaur, C., Vogt, J.L., Walker, M.A., Weherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL plate: 28 Row: 1 Column: 22.
Location/Qualifiers
1. 1908
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:20293 IMAGE:4121450"
/tissue_type="Muscle, thaladomysarcoma"

clone_11b="NIH_MGC_17"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1. 1908
/gene="OSGEP1"
/db_xref="locusID:64172"
104.1348
/gene="OSGEP1"
/product="OSGEP1 protein"
/codon_start=1
/protein_id="AAH1904.1"
/db_xref="GI:15080282"
/db_xref="locusID:64172"
/translation="MLILITKAGVFFPKRKYVEFLRSFNPHTLFLKIVLGIET
SCDDTAADVDETNGEALHSQTEVHLKKGIVPPAQQIHRNRIQIVQALSLAS
GSPSDLSAATIKPKGLALSLGVSFLQVQPKKPIPIHMEHALITLTK
VEPFLVILISGHCILALVQVSDPILIGKSLIDPDMDKVARRSLIKPECS
MSGKALEHLAKGNRHPFDIKPLHLKAKNDPSFTGQHTQTKIMKEKEBEG
QILSSADIAATVQHTACHLVKRTKRALIFCKPRDLIPNNAVLVASGVASNPYIR
FALEILTNATQCTLLCPPRCTDNGIMAVNGIERLRAAGILHIDIEGIRYBRCPL
GVDSKEVGEASIKVPLKMEI"
215.1111
/gene="OSGEP1"
/note="Peptidase M22; Region: Glycoprotease family"
/db_xref="CDD:pfam00814"

misc_feature
US-10-649-273-2_COPY_148_414 (1-267) x BC011904 (1-1908)
Alignment Scores:
Pred. No.: 1.97e-126 Length: 1908
Score: 1385.00 Matches: 267
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

ORIGIN
1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
545 ATGGAGGCTCAGCCTTACTTAATTAAGTTCACCAATTAAGTTCCTTTTACTT 604
21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
605 CTTTGGATTTCGAGAGTCACTGCTCTGTGGCATTAGTTCAGAGGTTTCAGATTTCG 664
41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 60
665 CTTCTGGAAGTCTTGGACATAGCACAGGTGACATGCTTGACAGGTGCAAGAAAGA 724
61 LeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyLysAlaIleGluHis 80
725 CTTCTTTTAATTAACATCCAGAGTCCACCAAGAGTGGGAAAGCCATTAGACAT 784
81 LeuAlaLysGlnGlyAsnArgPheHisAspLeuIleLysProProLeuHisHisAlaLys 100
785 TTGGCCAAACAGAAATGATTTTCAATTTTGAATCAACCTCCCTTGATATGCTAAA 844
101 AsnCysAspPheSerPheThrGlyLeuGlnHisValIleThrAspLysIleIleMetLysLys 120
845 AATTGTGATTTTCTTTTACTGACTCAACACGTTACTGATTAATAATAATGAAAAAG 904
121 GluLysGluGluGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaAla 140
905 GAAAAAGGAAAGGATTTGAGAAAGGCAAAATCTCTTCCAGACGACACTTCTGCC 964
141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
965 AAGATACGACACACATGATGATATCTTGAAAAAGAACATCGGCTATTTCTGTTT 1024
161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyVal 180
1025 TGTAAAGCAGAGAGCTTTTACTCTCAAAATAATGACAGTACTGTTGCATCTGCTGTGTC 1084

```

QY 181 AAserspNheTyrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThr 200
DB 1085 GGAAGTAACCTTCAATATCCGACAGACTCTGGAATTTTAAACCAACGACACAGTGCAC 1144
QY 201 LeuLeuCysProProProArgLeuCysThrAspAsnGlyIleMetIleAlaIleTyrAsnGly 220
DB 1145 TTGTTGTGTCCTCTCCAGACTATGACATGATTAAGGCAATTATGATGATGATGAAAGGT 1204
QY 221 IlleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGlu 240
DB 1205 ATTGAAAGACTACGTCTGCTGGCATTTTACATGACATTAAGGCAATCCGCTATGAA 1264
QY 241 ProLysCysProLeuGlyValAspIleSerLysGluValGlyGluAlaSerIleLysVal 260
DB 1265 CCAAAATGTCTCTTGAGATGACATATCAAAAGAAAGTTGGAGAGCTTCATTAAGTA 1324
QY 261 ProGlnLeuLysMetGluIle 267
DB 1325 CCACAATTAAATAAGAGATA 1345

RESULT 4
LOCUS AR428803 2197 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6642041.
ACCESSION AR428803
VERSION AR428803.1 GI:40188589
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2197)
AUTHORS Chen,J., Feder,J.N., Nelson,T.C., Krystek,S.R. and Duclos,F.
TITLE Polynucleotides encoding a novel metalloprotease, MP-1
JOURNAL Patent: US 6642041-A 1 04-NOV-2003;
FEATURES
source
1. 2197
location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 2,33e-126 length: 2197
Score: 1385.00 Matches: 267
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x AR428803 (1-2197)
QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
DB 672 ATGAGGCTCATGCACTTACTATTAGGTGACCAATTAAGTAGAATTTCTTTTATGTT 731
QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
DB 732 CTTTATATTTCTGAGAGTCACTGTCGTTGGCATTAAGTTCAGAGAGTTTCAATTTCTG 791
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
DB 792 CTTCTTGGAAGCTTTGGACATAGCACAGTGACATGCTTGACAAAGTGCGCAAGAGA 851
QY 61 LeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyLysAlaIleGluHis 80
DB 852 CTTTCTTTAATAAATCAATCAGAGTCTCCACATGATGATGGGAAAGCCATTAAGACAT 911
QY 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLys 100
DB 912 TTGGCCAAACAAAGAAATAGATTTCATTTTGACATCAAACTCCCTTGCAATCAGCTAAA 971
QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLys 120
DB 972 AATTGATTTTCTTTACTGACATTCAACACGTTTACTATTAATAATTAATCAAAAG 1031

```

```

QY 121 GluLysGluGluGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 140
DB 1032 GAAAAGAGGAGAGTATTGGAAGGCGCAATCTGCTTCAGCAGCAGACATGCTGCC 1091
QY 141 ThrValGlnHisGlyMetAlaCysHisLeuValLysArgThrHisAlaGalaIleLeuPhe 160
DB 1092 ACAGTACAGACACCAATGGCATGTCATCTGTGAAAAGAACACATCGGCTAATTCGTGT 1151
QY 161 CysLysGlnArgAspIleLeuLeuProGlnAsnAlaValIleuValAlaSerGlyVal 180
DB 1152 TGTAAAGAGAGAGCTGTGTAATCTCAAAATTAATACAGTACGTGTTGATTCGTGGATTC 1211
QY 181 AAserspNheTyrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThr 200
DB 1212 GCAAGTAACCTTCAATATCCGACAGCTCTGGAATTTTAAACAAACGACACAGTGCAC 1271
QY 201 LeuLeuCysProProProArgLeuCysThrAspAsnGlyIleMetIleAlaIleTyrAsnGly 220
DB 1272 TTGTTGTGTCCTCTCCAGACTATGACATGATTAAGGCAATTATGATGATGATGATGAT 1331
QY 221 IlleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGlu 240
DB 1332 ATTGAAAGACTACGTCTGCTGGCATTTTACATGACATTAAGGCAATCCGCTATGAA 1391
QY 241 ProLysCysProLeuGlyValAspIleSerLysGluValGlyGluAlaSerIleLysVal 260
DB 1392 CCAAAATGTCTCTTGAGATGACATATCAAAAGAAAGTTGGAGAGCTTCATTAAGTA 1451
QY 261 ProGlnLeuLysMetGluIle 267
DB 1452 CCACAATTAAATAAGAGATA 1472

RESULT 5
LOCUS AR428808 1387 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 21 from patent US 6642041.
ACCESSION AR428808
VERSION AR428808.1 GI:40188594
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1387)
AUTHORS Chen,J., Feder,J.N., Nelson,T.C., Krystek,S.R. and Duclos,F.
TITLE Polynucleotides encoding a novel metalloprotease, MP-1
JOURNAL Patent: US 6642041-A 21 04-NOV-2003;
FEATURES
source
1. 1387
location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 5,43e-124 length: 1387
Score: 1358.50 Matches: 266
Percent Similarity: 91.10% Conservative: 0
Best Local Similarity: 91.10% Mismatches: 1
Query Match: 98.09% Indels: 25
Gaps: 1

US-10-649-273-2_COPY_148_414 (1-267) x AR428808 (1-1387)
QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
DB 465 ATGAGGCTCATGCACTTACTATTAGGTGACCAATTAAGTAGAATTTCTTTTATGTT 524
QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
DB 525 CTTTATATTTCTGAGAGTCACTGTCGTTGGCATTAAGTTCAGAGAGTTTCAATTTCTG 584
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60

```

DB 585 CTTCTTGAAGAGCTTTGGACATAGCACAGGTGACATGCTTGAACAAGTGGCAAGAGA 644
QY LeuSerLeuIleuYshisProGluCySerTherMetSerGlyLyalaIleGluHis 80
DB 645 CTTTCTTTAATAAACAATCCAGAGTGTCTCCACCAATGAGTGTGGAAAGCATAGAGCAT 704
QY LeuAlaIysGlnGlyAsnArgPheHisPheAspIleuYsProPoleuHisAlaIlys 100
DB 705 TTGGCCAAACAGAGAAATAGATTTCATTGTGACATCAAACTCCCTTGACATAGCTAAA 764
QY 101 AsnCyAspPheSerPheThnGlyLeuGlnHisValThrAspIysIleIleMetIysIys 120
DB 765 AATTGTGATTTTCTTTTACTGACCTTCAACACGTTACTGATTAATAATATGAAAAAG 824
QY 121 GluIysGlnGlyIle-----GluYsGlyGlnIleLeuSerSerAla 126
DB 825 GAAAAAGGAGAGATATTTCTTAATTAAGTAAGTGAACGATTAATATTCCTGGATTG 884
QY 127 -----GluYsGlyGlnIleLeuSerSerAla 135
DB 885 TGCCATAAATAAGTGTCTCATTTCTGACAGTATGAGAGGGCAATCTCTCTTCAACA 944
QY 136 AlaAspIleAlaIleThrValGlnHisTherMetAlaCyshisIleuValIysArgThrHis 155
DB 945 GCAGACATTTGCTGTCACAGTACAGACACAATGGCATGTCTTGTGAAAAAGAACACAT 1004
QY 156 ArgAlaIleLeuPheCyIysGlnArgAspIleuProGlnAsnAsnAlaValIleuVal 175
DB 1005 CGGGCTATTCTGTTTGTGAACAGAGACTGTGTACTCTCAAAATTAAGGAGTCTGTT 1064
QY 176 AlaSerGlyGlyValAlaSerAsnPheTyrlleArgArgAlaIleGluIleLeuThrAsn 195
DB 1065 GCATCTGTGTGTGTGCGAAGTACTTCTATATCCGACAGCTCTGGAATTTTAAACAAC 1124
QY 196 AlaThrGlnCyThrLeuLeuCyAspProProArgIleuCyGlnAspAspGlyIleMet 215
DB 1125 GCAACACAGTGCATTTGTGTGTCTCTCCACAGACTATGACATGATTAATGGCATTAAG 1184
QY 216 IleAlaIleArgGlyIleGluArgLeuArgAlaGlyIleGluIleLeuHisAspIleGlu 235
DB 1185 ATTGCAATGAGATGTATTTGAAGAACTACGTGTGGCTTGGCATTTTACATGACATGAA 1244
QY 236 GlyIleArgGlyArgProIysCyAspProLeuGlyValAspIleSerGlyValIleGlu 255
DB 1245 GGCATCCGCTATGAAACCAAAATGTCTCTTGAAGTACATATCAAAAGAAAGTGGAGAA 1304
QY 256 AlaSerIleIysValProGlnLeuIleMetGluIle 267
DB 1305 CTTTCCATAAAGTACCAACATTAATAAATGAGAGATA 1340
RESULT 6
LOCUS HSA295148 1387 bp mRNA linear PRI 30-OCT-2000
DEFINITION Homo sapiens mRNA for putative sialoglycoprotease type 2.
ACCESSION AJ295148
VERSION AJ295148.1 GI:11071726
KEYWORDS metallopeptidase; sialoglycoprotease.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Chen, J.M., Fortunato, M. and Barrett, A.J.
TITLE Cloning and sequencing of a second human putative
sialoglycoprotease homologue
JOURNAL Unpublished
AUTHORS Chen, J.M.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-2000) Chen J.M., MRC Molecular Enzymology
Laboratory, The Babraham Institute, Babraham, Cambridge, CB2 4AT,
UNITED KINGDOM
FEATURES Location/Qualifiers

SOURCE 1. 1387
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue="uterus"
CDS
24. 1343
/function="metallopeptidase of family M22"
/codon_start=1
/product="putative sialoglycoprotease type 2"
/protein_id="CAC14666.1"
/db_xref="GI:11071727"
/db_xref="GOA:O9H4B0"
/db_xref="UniProt/TREMBL:O9H4B0"
/translation="MLILTKTGVFPEKSKRVVEFLSFFNPETLFLKILVIGLET
SCDDTAADVDEGTAVLEAIHSQTEVHLTKGVIVPPAQQLHRENIQRIYQVLAISAS
GVSPDSIAIVTTIKRGLALSLGVLSFSLQVQKKPFIPHHMAHALITRLTNK
VEPPIVLTISGGHCLALVGVSPFLIGSLIDLPDMDLKVRLSLIKHPCST
MSGKALIEHLAKQGRFHPDIKPLPHAKNCDPSTGLOHTYDILMKKEBEGFLI
SKVQINIPGLCLKIAAHFCREKQILSSADIAVQHTACHLVKTRHAILFCK
ORDLLPQNNAVLVASGVASNFYIRRALEILTNAQCTLCPPLCTDNGIMIAMNG
IERLRGIGILHIDIEGIRYBPCKPLGVDISKEVBASIKVPLKMEI"
ORIGIN
Alignment Scores:
Pred. No.: 5,43e-124 Length: 1387
Score: 1358.50 Matches: 266
Percent Similarity: 91.10% Conservative: 0
Best Local Similarity: 91.10% Mismatches: 1
Query Match: 98.09% Indels: 25
DB: 9 Gaps: 1
US-10-649-273-2_copy_148_414 (1-267) x HSA295148 (1-1387)
QY 1 MetGluAlaHisAlaLeuThrIleArgLeuLeuThrAsnIysValGluPheProPheLeuVal 20
DB 465 ATGAGGCTCATGACCTTACTATAGTTGACCAATAAGTGAATTCCTTTTATGTT 524
QY 21 LeuLeuIleSerGlyGlyHisCySerLeuAlaIleValGlnGlyValSerAspPheLeu 40
DB 525 CTTTGAATTTCTGAGGTCACGTCCTGTGGCATTAAGTTCAAGAGCTTCAATTTTCTG 584
QY 41 LeuLeuGlyIysSerLeuAspIleAlaProGlyAspMetLeuAspIysValAlaArgArg 60
DB 585 CTTCTTGAAGAGCTTTTGGACATACACAGGTGACATGCTTGACAGAGTGGCAAGAGA 644
QY 61 LeuSerLeuIleYshisProGluCySerTherMetSerGlyLyalaIleGluHis 80
DB 645 CTTTCTTTAATAAACAATCCAGAGTGTCTCCACCAATGAGTGTGGAAAGCATAGAGCAT 704
QY 81 LeuAlaIysGlnGlyAsnArgPheHisPheAspIleuYsProPoleuHisAlaIlys 100
DB 705 TTGGCCAAACAGAGAAATAGATTTCATTGTGACATCAAACTCCCTTGACATAGCTAAA 764
QY 101 AsnCyAspPheSerPheThnGlyLeuGlnHisValThrAspIysIleIleMetIysIys 120
DB 765 AATTGTGATTTTCTTTTACTGACCTTCAACACGTTACTGATTAATAATATGAAAAAG 824
QY 121 GluIysGlnGlyIle-----GluYsGlyGlnIleLeuSerSerAla 126
DB 825 GAAAAAGGAGAGATATTTCTTAATTAAGTGAACGATTAATATTCCTGGATTG 884
QY 127 -----GluYsGlyGlnIleLeuSerSerAla 135
DB 885 TGCCATAAATAAGTGTCTCATTTCTGACAGTATGAGAGGGCAATCTCTCTTCAACA 944
QY 136 AlaAspIleAlaIleThrValGlnHisTherMetAlaCyshisIleuValIysArgThrHis 155
DB 945 GCAGACATTTGCTGTCACAGTACAGACACAATGGCATGTCTTGTGAAAAAGAACACAT 1004
QY 156 ArgAlaIleLeuPheCyIysGlnArgAspIleuProGlnAsnAsnAlaValIleuVal 175
DB 1005 CGGGCTATTCTGTTTGTGAACAGAGACTGTGTACTCTCAAAATTAAGGAGTCTGTT 1064

QY 176 AlaSerGlyValAlaSerAspPheTyrIleArgArgAlaLeuGluIleLeuThrAsn 195
| | | | |
Db 1065 GCATGCGTGTGTCCAGAGTAACCTTATATCCGAGAGCTTCGAAATTTAAACAAC 1124
| | | | |
QY 196 AlaThrGlnCysThrLeuLeuGlyProProArgLeuGlySerThrAspAsnGlyIleMet 215
| | | | |
Db 1125 GCAACACAGTGCACCTTGTGTGTCTCCCTCCAGACTTACGACTGATTAATGCGATTATG 1184
| | | | |
QY 216 IleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGlu 235
| | | | |
Db 1185 ATTCGATGGAATGCTATTGGAAGACATACGTCGTGGCATTTACATGACATGAA 1244
| | | | |
QY 236 GlyIleArgTyrGluProLysCysProLeuGlyValAspIleSerGlyValGlyGlu 255
| | | | |
Db 1245 GGCATCCGCTATGACCAAAATGTCCTCTGGAGTACATATCAAAAGAAAGTTGAGAA 1304
| | | | |
QY 256 AlaSerIleLeuValProGlnLeuLysMetGluIle 267
| | | | |
Db 1305 GCTTCATAAAGTACCACAAATTAAATGAGATA 1340
| | | | |
RESULT 7
AX664697 1245 bp DNA linear PAT 22-MAR-2003
LOCUS Sequence 6 from Patent WO02074960.
ACCESSION AX664697
VERSION AX664697.1 GI:29164457
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1
TITLE Leiby, K.R., Kapeller-Libermann, R. and Glucksmann, M.
JOURNAL 38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions
of human proteins and uses thereof
Patent: WO 02074960-A 6 26-SEP-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source 1.1245
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 5.14e-123 Length: 1245
Score: 1348.00 Matches: 260
Percent Similarity: 98.50% Conservative: 3
Best Local Similarity: 97.38% Mismatches: 4
Query Match: 97.33% Indels: 0
Gaps: 0
DB: 6
US-10-649-273-2_COPY_148_414 (1-267) x AX664697 (1-1245)
QY 1 MetGluAlaHisIleAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
| | | | |
Db 442 ATGGAGGCTCAATGCACTTACTATTAGGTTACCAATAAGTAGAATTTCTTTTATGTT 501
| | | | |
QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValaSerAspPheLeu 40
| | | | |
Db 502 CTTTCATTTCTTGAGAGTCACTGCTGTGGCATTTAGTCAAGAGTTTCAGATTTTCTG 561
| | | | |
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValaIleArgArg 60
| | | | |
Db 562 CTTCTTGAGAAAGCTTTTGACATAGCACCAGGTGACATGCTTGACAGAGTGGCAAGAGA 621
| | | | |
QY 61 LeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyLysValaIleGluHis 80
| | | | |
Db 622 CTTTCATTTAAATAAACAATCCAGAGTCTCCACCAATGAGTGTGGAGAACCATAGAACAT 681
| | | | |
QY 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisIleAlaLys 100
| | | | |
Db 682 TTGGCCAAACAAGGAATATGATTTTCATTTTTCATGATCAACCTCCTTGCATCATGCTAAA 741
| | | | |

QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
| | | | |
Db 742 AATGTGATTTTCTTTTATCGAGCTTCAACACGCTTACTGCTAATAAATATGAAACACAG 801
| | | | |
QY 121 GluLysGlnGluGlyIleGlyLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaIle 140
| | | | |
Db 802 AAACAAAGAGAAAGGATATGGAAGGGGCAATCTGCTTTCAGACAGACATTCGTCGC 861
| | | | |
QY 141 ThrValGlnHisIleThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
| | | | |
Db 862 ACACTACAGACACAAATGGCATGTCTATCTGTGAAAAGAAACATCGGGCATTTCTGTT 921
| | | | |
QY 161 CysLysGlnArgAspLeuLeuProGlnAsnAlaValLeuValaIleSerGlyValaIle 180
| | | | |
Db 922 TGTATGACAGAGAGACTTGTATACCTCAAAATATATGACAGTACTGTTGATCTGGTGTCT 981
| | | | |
QY 181 AlaSerAsnPheTyrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThr 200
| | | | |
Db 982 GCAAGTAACTTCTATATCCGAGAGCTCTGGAATTTTAAACAAACGCAACACAGTGCCT 1041
| | | | |
QY 201 LeuLeuGlyProProArgLeuGlySerThrAspAsnGlyIleMetIleAlaTrpAsnGly 220
| | | | |
Db 1042 TTGTGTGTCCTCTCCAGACTATGACATGATATGACATTATGATGATGAGATGCT 1101
| | | | |
QY 221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGlu 240
| | | | |
Db 1102 ATTCGAAAGACTACCTGCTGCGCTTGGCATTTTACATGACATAGAAAGCATCGCATGAA 1161
| | | | |
QY 241 ProLysCysProLeuGlyValaAspIleSerGlyValaGlyGluAlaSerIleLysVal 260
| | | | |
Db 1162 CCAAAATGTCCTCTTGAGATGACATATCAAAAGAAAGTTGAGAGAGCTTCATTAAGATA 1221
| | | | |
QY 261 ProGlnLeuLysMetGluIle 267
| | | | |
Db 1222 CCACAAATTAATAAATGAGAGATA 1242
| | | | |
RESULT 8
AX664695 1820 bp DNA linear PAT 22-MAR-2003
LOCUS Sequence 4 from Patent WO02074960.
ACCESSION AX664695
VERSION AX664695.1 GI:29164455
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1
TITLE Leiby, K.R., Kapeller-Libermann, R. and Glucksmann, M.
JOURNAL 38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions
of human proteins and uses thereof
Patent: WO 02074960-A 4 26-SEP-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source 1.1820
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD80044.1"
/db_xref="GI:29164456"
/translation="MLILRTKAGVFFKRSKKRYEFLRSFNFHGTLLFLKILVIGIET
SCDDTAADVDETNVGEBAIHSQTEVHLTKGIVPPAAQQLRENTQRIYQVBLAS
GVSPSDSAVATTIKPGALSIGVLSFSLQVQLAKKPFPIHMHMAHLITRLTK
VEPPTVLISGGHCILAIQVQSDPFLIKSLIDIAQDMIDKVARRLSLIKPESCT
MSGKAIEMHAKQGNRFPDIKPLIHAKCDSFPTGLQHTTDQNERKRBGIEKG
QILSSAADIAVQHTVACHLVKRTTHAIIIFKQRDLIPQNNAVLVASGVASNFYLR
RALEILTNATCTLLCPRLCTDNGIMAWNGIERLRAGIGIILHDIIGIIEBKPL
GVDSKEVGASIKVPLKMEI"
ORIGIN

Alignment Scores:

Pred. No.:	8, 17e-123	Length:	1820
Score:	1348.00	Matches:	260
Percent Similarity:	98.50%	Conservative:	3
Best Local Similarity:	97.38%	Mismatches:	4
Query Match:	97.33%	Indels:	0
DB:	6	Gaps:	0

US-10-649-273-2_COPY_148_414 (1-267) x AK666695 (1-1820)

```

OY      1 MetGluAlaHisaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
DB      587 ATGAGAGGCTCATGCACTTACTATTAGGTGACCAATTAAGTAGAATTTCTTTTATTGTT 646
OY      21 LeuLeuIleSerCylYglYHisCysLeuLeuAlaLeuValGlnGlyValIleSerAspPheLeu 40
DB      647 CTTTGTGATTTCTGGAGGTCACTGTCTGTGGCATTTAGTTCAAGGAGTTTCAGATTTCCTG 706
OY      41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
DB      707 CTTCTTGGAAAGCTTTTGGACATAGCACCGGTGACATGCTTTCAGCAAGTGGCAAGAGA 766
OY      61 LeuSerLeuIleLysHisProGlyCysSerThrMetSerGlyGlyValAlaIleGluHis 80
DB      767 CTTTCTTATATAAATCATCCAGAGTGTCTCCACCATGATGCTGGGAAGCATAGAACAT 826
OY      81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProPheLeuHisAlaLys 100
DB      827 TTGGCCAAACAAAGAAATGATTTCAATTTTGAACATCAACCTCTTGCATCATGCTTAA 886
OY      101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
DB      887 AATTGATTTTCTTTTACTGACATTCACACGTTACTGATTAATAAATGAACACAG 946
OY      121 GluLysGlnGlyLysIleGlyLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 140
DB      947 AAACAAAGAGAGAGATTGAGAAAGGGCAATCTGTCTTCAGACAGACATTCGTGCC 1006
OY      141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
DB      1007 AAGTACAGACACAAATGCGATGTCATCTGTGAAGAAACACATCGGGCTATTCTGTT 1066
OY      161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyVal 180
DB      1067 TGTACACAGAGACATTTGACTCTCAAAATATGACATGCTGTCATCTGGTGTGTC 1126
OY      181 AlaSerAspPheThrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr 200
DB      1127 GCAAGTAACTTCTATATCCGAGAGCTCTGAAATTTTACAAACGCAACACAGTCACT 1186
OY      201 LeuLeuCysProProPheArgLeuCysThrAspAsnGlyIleMetIleAlaTPAsnGly 220
DB      1187 TTCTTGTCTCTCTCTCCACAGTATGACATGATATGACATTAATGATTCATGAGATGCT 1246
OY      221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGlu 240
DB      1247 ATGGAAGAAGTACGTCGTGCTGGCATTTTATCATGACATAGAAAGCATCCGCTATCAA 1306
OY      241 ProLysCysProLeuGlyValAspIleSerLysGlyValGlyGlyAlaSerIleLysVal 260
DB      1307 CCAAAATGTCCTCTTGGAGTAGACATATCAAAAGAGTTGAGAGCTTCATTAAGAATA 1366
OY      261 ProGlnLeuLysMetGluIle 267
DB      1367 CCAACAATTAATAAAGAGATA 1387

```

RESULT 9
BC058172 1844 bp mRNA linear ROD 08-OCT-2003
DEFINITION Mus musculus cDNA clone MGC:67870 IMAGE:5012054, complete cds.
ACCESSION BC058172
VERSION BC058172.1 GI:34849663

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

MGC.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1844)

Strausberg, R.D., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenman, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Schaefer, T.E., Brownstein, M.J., Udell, T.B., Toshitsuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

Mckernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalins, D.E.,

Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 1844)

Strausberg, R.

Direct Submission

Submitted (15-SEP-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: GGC help desk

Email: ggc-help@nih.govWeb site: <http://www.nisc.nih.gov/>Contact: nisc_mgc@nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooker, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,

Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, V.C.,

McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W.,

Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/BLM at: <http://image.llnl.gov>

Series: IRAC Plate: 123 Row: d Column: 22

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 21312463.

Location/Qualifiers

1..1844

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CZECH IT"

/db_xref="taxon:10090"

/clone="MGC:67870 IMAGE:5012054"

/tissue_type="mammary tumor metastasized to lung. Tumor

arose spontaneously from a senescent normal mammary

(clonal) outgrowth infected with the virus MMTV."

/clone_lib="NCI_CGAP_Lu29"

/lab_host="DH10B"

CDS

222..1466 /note="Vector: PCMV-SPORT6"

/codon_start=1

/product="unknown (protein for MGC:67870)"

/protein_id="AAH58172.1"

/db_xref="GI:34849664"

/translation="MLMERTAGAIKPKPKSKYKYLRRSPVPTLSCHKVLGIEIT
SCDDTAAVAVDETGNVLEALHSOTVHLEGTGIVPPVAQQLRENIQRIVEETLSAS
RTTPDLISAATTIKPGLASLGVLSFSLQVNRPKKPIPIHMEAHALITRLTKK
VEFPLVLISGGHCLALVQGVDFLLGKSLIDIAQMDLKVARSLSIKPECSY
MSGKALIEHLAKODGNFHTINPMDNAKCDPSFTGLQHTIKLTKHKEGSIENG
OLSSADIAAANOAHATACHLAKRTERRALIFCKOKLSPANVLVSVGGSEINLYR
KALEIVANAQCTCLCPPEPLCTDNGIMIANNGIEXLRAGLGVLDVEDIRYKPKCL
GIDISREVAEAAIKVPLKXAL"

misc_feature

333..1397 /note="OR17: Region: Metal-dependent proteases with
possible chaperone activity [posttranslational
modification, protein turnover, chaperones]"

/db_xref="CDD:COG0533"

ORIGIN

Alignment Scores:

Pred. No.:	2,76e-111	length:	1844
Score:	1231.00	Matches:	233
Percent Similarity:	92.88%	Conservative:	15
Best Local Similarity:	87.27%	Mismatches:	19
Query Match:	88.88%	Indels:	0
DB:	10	Gaps:	0

US-10-649-273-2_COPY_148_414 (1-267) x BC058172 (1-1844)

1 MetGUALAHISALALEUTHRIIEARGLEUTHRASNLVSVALGIUPHEPROHELEUVAL 20
663 ATGGAGGCTCAGCAGCTGATTAAGGCTCACCAATTAAGATTTCTTTCTTAGTT 722
21 LeuLeuIISeSerGIyGIyGIyHISySLeuLeuAlaLeuValGIyGIyValSeArSpHeLeu 40
723 CTTTTCATTTCTTGCGCGTCACTCCCTGTTGGCATTAGTCCAGGCTGTTCCGATTTCTCG 782
41 LeuLeuGIyLySeSerLeuAspIIeAlaProGIyAspMetLeuAspLySValAlaArgArg 60
783 CTCCTTGGGAAGCTTTGACATAGCGCCAGCGGACATGCTTGAACAAGGTGGCAAGAGA 842
61 LeuSeSerLeuIIeLySHISProGIyCySeSerThMetSeSerGIyGIyValaIIeGIyHIS 80
843 CTTTCTTATCAACATCCAGATGTTCTTACATAGTGTGGAAAGCTATAGAACAT 902
81 LeuAlaLySeGIyGIyAsnArgPheHISpHeAspIIeLySProProLeuHISHSALaLyS 100
903 TTGGCCAAAGACCGGAATAGATTCATTTTACTATCAATCCACTATGACAGAATGCTAAG 962
101 AsnCyAspPheSeSerPheThrGIyLeuGIyHISValaThrAspLySIIeIIeMetLyLyS 120
963 AATTGGGATTTTCTTTTACGGGACCTTCAACATATTAAGTAATTAACACACACACAG 1022
121 GIyLySGIyGIyGIyIIeGIyLySGIyGIyIIeLeuSeSerAlaAlaAspIIeAlaAla 140
1023 GAAAAAGAAAGAGGCTTGAAGGGGCAATCTCTGTCACTCACTGACAGCATTTGCTGCT 1082
141 ThrValGIyHISThrMetAlaCySHISLeuValLySArgThrHISArgAlaIIeLeuPhe 160
1083 GGGGTACAGCATCAACAGGCTGCCACTTGCAAAAAGAACACATGCTGCTATCTGTTT 1142
161 CyLySeGIyHISArgAspLeuLeuProGIyAsnAsnAlaValaLeuValaAlaSeSerGIyVal 180
1143 TCCAAAGCAAAAAATTTGCTATCTCCAGCTTACAGATATTAAGTTGATCTGAGGGGTGT 1202
181 AlaseArSpPheThrIIEARGARGAlaLeuGIyIIeLeuThrAsnAlaThrGIyCyThr 200
1203 GGAAGTACTGTGATCCGAAAAGCATGGAATGTGCCAATATGCAAGCGAGTGCAGC 1262
201 LeuLeuCySeProProProArgLeuCyThrAspAsnGIyIIeMetIIeAlaTPAsnGIy 220

Db 1263 TTGCTGTGTCCTCCCTCCAAAGACTGTGACATGACATGACATGATTCATGCAATGCA 1322

Qy 221 IIeGIyArgLeuArgAlaGIyLeuGIyIIeLeuHISAspIIeGIyGIyIIEArgYrGIu 240

Db 1323 ATTCGAAAGATTAAGTGTGCTGGCTGGCTTTTATCATATGATGAGAACATCGATATGAA 1382

Qy 241 ProLyCySeProLeuGIyValaAspIIeSerLySeGIyValaGIyGIyIIeAlaSeIIeLySVal 260

Db 1383 CCAAAAGTCTCTTGGAAATGACATATTCAGAGAGGTGGCAAGCTGCCATTAAGTA 1442

Qy 261 ProGIyLeuLySmetGIuIIe 267

Db 1443 CCGGATTAATAATGCACTT 1463

RESULT 10

BC038910

LOCUS

DEFINITION Mus musculus O-6-ialaloycoprotein endopeptidase-like 1, mRNA (cDNA

ACCESSION BC038910

VERSION BC038910

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Mus musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 1017)

Strausberg R.D., Collins F.S., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullihy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E., Jones S.J. and Marz M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL MEDLINE 22388257

PUBMED 12477932

REFERENCE

2 (bases 1 to 1017)

Strausberg R.

Direct Submission

Submitted (25-OCT-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: Jeffrey B. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

BC Cancer Agency, Vancouver, BC, Canada

info@cgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Diana Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stolt, Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy,

George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRK Plate: 86 Row: F Column: 12.

FEATURES

source

1. 1017

/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5053559"
/issue_type="Liver, normal, 5 month old male mouse."
/clone_lib="NCI CGAP_L19"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:

Pred. No.:	6,53e-111	Length:	1017
Score:	1224.00	Matches:	233
Percent Similarity:	92.51%	Conservative:	14
Best Local Similarity:	87.27%	Mismatches:	20
Query Match:	88.38%	Indels:	0
DB:	10	Gaps:	0

US-10-649-273-2_COPY_148_414 (1-267) x BC038910 (1-1017)

1 MetGUAHHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
26 ATGAGAGGCTCAGCAGCTACTATTAGGCTCACCAATTAAGTAGAATTTCTTTTGT 85
21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
86 CTTTATATTCTGGCGGTCACTGCTGTTGGCATTAGTCCAGAGTGTTCGATTTCCTG 145
41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
146 CTCCTTGGAGAGCTTTGGACATAGCACAGGCGACATGCTTACAGAGTGGCAGAGAGA 205
61 LeuSerLeuIleLysHisProGlyCysSerThrMetSerGlyGlyValAlaIleGlnHis 80
206 CTTTCTTATATCAACATCCAGATGTTCTACATGATGAGTGGAGAAAGCATATGAAAG 265
81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLys 100
266 TTGGCCAAAGACCGAATATGATTCATTACTATCAATCACTACCTAGCAGAAAGCTAG 325
101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
326 AATTGGATTCTTTCTTTCACAGGAGCTTCACATATTACTATTAAGCTTAATTAACACAC 385
121 GluLysGlnGluGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaAla 140
386 GAAAAAGAAAGAGGCTTGGAGAGGCGCAATCTGTGATCACTGACGACATTTGCTGCT 445
141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
446 GCGGTACAGATCCACAGCGTGCACCTTGCAGAAAGAACACATCCGCTATTTCTGTT 505
161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyVal 180
506 TCGAACAGAAAAATTGCTCTCTCCAGCTAACCCAGATATTAGTGTATGAGAGGTGTT 565
181 AlaSerAsnPheThrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr 200
566 GCAAGTAATCTTGACATCCAAAGACATGGAAATGTGCGAAATGCAACCCAGTGCAGC 625
201 LeuLeuCysProProArgLeuCysThrAspAsnGlyIleMetIleAlaTTPAsnGly 220
626 TTCTTGTGTCACCTCCAAAGACTGTGACATGACATGACATGATTCATGATGAAATGGA 685
221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGlu 240

|||||
686 ATTAGAAGATTACGCTGCGCTTGGCTTTTACATGATGAGAGACATCCGATATGAA 745

241 ProLysCysProLeuGlyValAspIleSerLysGluValGlyGlnAlaSerIleLysVal 260
746 CCAAAATGCTCTTGGAGTGAATGACATATCCAGAGAGTTGCAAGAGCTGCCATTAAGTA 805

261 ProGlnLeuLysMetGluIle 267
806 CCGCATTTAAAAATGCGACTT 826

RESULT 11

AX713716 2208 bp DNA linear PAT 15-APR-2003
LOCUS AX713716
DEFINITION Sequence 400 from Patent EP1293569.
ACCESSION AX713716
VERSION AX713716.1 GI:29888642
KEYWORDS

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
Masuhio,Y.

TITLE

Patent: EP 1293569-A 400 19-MAR-2003.

JOURNAL

Helix Research Institute (JP) ; Research Association for

FEATURES

Biotechnology (JP)

source

1. 2208
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	1.57e-108	Length:	2208
Score:	1204.00	Matches:	239
Percent Similarity:	89.51%	Conservative:	0
Best Local Similarity:	89.51%	Mismatches:	4
Query Match:	86.93%	Indels:	24
DB:	6	Gaps:	1

US-10-649-273-2_COPY_148_414 (1-267) x AX713716 (1-2208)

1 MetGUAHHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
785 ATGAGAGGCTCAGCAGCTACTATTAGGCTCACCAATTAAGTAGAATTTCTTTTGT 844
21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
845 CTTTGTGATTCTGAGAGCTGCTGTGCTGTCATTAGTTCAGAGGATTTCTG 904
41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
905 CTTCTGGAAGGCTTTGGACATAGCACAGGTGACATGCTTGAACAAGTGGCAAGAGA 964
61 LeuSerLeuIleLysHisProGlyCysSerThrMetSerGlyGlyValAlaIleGlnHis 80
965 CTTCTTAAATTAACATCCAGAGCTCCACCAATGAGTGGAGAAAGCCATTAACACT 1024
81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLys 100
1025 TTGGCCAAAGAAAGAAATGATTTCACTTTTACATCAAACTCCTTGCATCATCTAAA 1084
101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
1085 AATTGTGATTCTTCTTATCTGAGCTTCAACCGTTACATGAATTAATATATGAAAAAG 1144
121 GluLysGlnGluGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaAla 140

```

Db      1145 GAAAAAGAGGAGGATGTTGAGAGGGGCAATCTGCTTCAGAGAGCAATTCGCGCC
Oy      141  TTTTAAIGTnhiThrmctalaCyShiEleuVallysaRgThriSaRgaIailelePhe
Db      1205 ACGGTACAGACACATGATGATCTTGTGAAAAAGAACACATCGGGCTATTCGTTT
Oy      161  CyelYsglnaRgaRpleuLeuProglInaSaMaAlaValleuValaIaaserglyVal
Db      1265 TGTAAACAGAGAGACTTTGTTACCTCAAAATATATGCACTAGCTGTCATCGTGGTGC
Oy      181  AlaserAnpHeYrTleAgaRgaIaaleuGlulileuThraMaIathRlnCyEthr
Db      1325 GCAAGTAACTCTGATATCCGACAGCTCTGAAATTTTAAACAAACGACACAGTGCAC
Oy      201  LeuLeuCyEthrProProAaRgaLeuCyEThraSpaMnglyIleMetIleAlaIraPaMngly
Db      1385 TTGTTGTCTCTCTCCACACATGATGATGATGATGATGATGATGATGATGATGATGATG
Oy      221  IlegluArgLeuAgaIaaglyLeuGlyIleleuHisaPliIlegluGlyIleArgTyrGlu
Db      1435 -----
Oy      241  ProlyCyEProleuGlyValaPliIleSerlyGluValGlyIleValaIleValaIleVala
Db      1436 ---TGATGTCCTCTTGAGAGTGAATATCAAAAGAGTGGAGAGCTTCATCAAAAGTA
Oy      261  ProglLeuLysMetGluile 267
Db      1493 CCACAAATTAATAATGAGAGATA 1513

RESULT 12
AK055441
LOCUS      2208 bp mRNA linear PRI 30-JAN-2004
DEFINITION Homo sapiens cDNA FLJ30879 fis, clone FEBR2004552, highly similar
            to Homo sapiens mRNA for putative statoglycoprotease type 2.
ACCESSION AK055441
VERSION    AK055441.1 GI:16550166
KEYWORDS   oligo capping; f1s (full insert sequence).
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS    Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
            Makamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
            Sekine, M., Oobayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
            Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
            Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
            Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
            Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Puruya, T.,
            Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
            Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
            Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
            Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
            Yoshida, M., Hattori, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
            Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, P., Hara, R.,
            Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuki, H., Oshima, A.,
            Sasaki, N., Aotsuka, S., Yoshioka, Y., Matsunawa, H., Ichihara, T.,
            Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
            Teraishina, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizouchi, H.,
            Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
            Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
            Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
            Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
            Ozaki, K., Hiro, M., Omori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
            Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Matsunura, K.,
            Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsunaga, K.,
            Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,
            Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
            Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
            Nagase, T., Nomura, N., Kikuchi, H., Masuno, Y., Yamashita, R.,
            Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.

```

```

TITLE      Complete sequencing and characterization of 21,243 full-length
JOURNAL    human cDNAs
NAT. GENET. 36 (1), 40-45 (2004)
PUBMED     14702039
REFERENCE
AUTHORS    Nishi, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,
            Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Makamatsu, A., Ishii, S.,
            Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
            Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
            Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,
            Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
            Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.
TITLE      Unpublished
JOURNAL    3 (bases 1 to 2208)
REFERENCE  Isogai, T., Otsuki, T. and Sugiyama, T.
AUTHORS    Direct Submission
TITLE      Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail: genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT    NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: RAB and
            HRI.
FEATURES
SOURCE
LOCATION/Qualifiers
1..2208
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FEBR2004592"
/tissue_type="brain"
/clone_id="FEBR2"
/dev_stage="fetus"
/note="cloning vector: pME18SFL3"
344..1438
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB70923.1"
/db_xref="GI:16550167"
/translation="MTLITKTAGVFFPKSKRVYELRSFNPFGTLFLKRLVIGLET
SCDPTAAVNDENGTGEAHSOTVEHLXGVIYPAQOULHRENIORIVOEALISAS
GVSPSDSAITTKTPGALSUGVSGSLQVLQKRPFIHHEMNAALITRLTK
VEPPLVLVLSGGCLALVGVSDFLVGLKSLDLPDMDLVKRRRLPLIKPCEST
MSGKALIEHLAKQNRFFIDIKPPLHAKNCDPFTGLQHVYDKIIMKKEKEGIEG
QILSSADIAITVQHTWACHLVKTRHAILFLCKQPDLLPQNNAVLVASGVASNFCIR
RALEILTNAGCTILCPPPRICDINGMIA"
ORIGIN
Alignment Scores:
Pred. No.: 1.57e-108 Length: 2208
Score: 1204.00 Matches: 239
Percent Similarity: 89.51% Conservative: 0
Best Local Similarity: 89.51% Mismatches: 4
Query Match: 96.93% Indels: 24
Gaps: 1
US-10-649-273-2_COPY_148_414 (1-267) x AK055441 (1-2208)
Oy      1  MetGUlaHISAlaLeuThrlleAryLeuThraMnlyValaGluPheProPheLeuVal 20
Db      785  ATGAGGCTCATGCACTTACATTAAGTGTGACCAATTAAGAAATTCCTTTTAACTT 844
Oy      21  LeuLeuIleSerglyGlyHISCyEthrLeuLeuAlaValaGlnGlyValaSerAspPheLeu 40
Db      845  CTTTGAATTTCTTGAGAGTCACTGCTGTGGCATTAAGTTCACAGAGCTTCAAGATTTTCG 904
Oy      41  LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValaIaArgArg 60
Db      905  CTTCTTGAAAGATCTTTTGACATAGCACACAGGTGACATGCTTGACACAGGTGCAAGAA 964

```

```

OY 61 LeuSerLeuIleIleYshisProGluCySerThrMetSerGlyGlyValAlaIleGluHis 80
DB 965 CTTCCCTTATATAAATCATCCAGAGTGTCTCACCACATGAGTGTTGGGAAGCATGAAACAT 1024
OY 81 LeuAlaIleGlnGlyValanArgPheHisPheAspIleLeuSerProPoleuHisHisAlaLys 100
DB 1025 TTGGCCAAACAAAGGAATATGATTTCAATTTTGACATCAAACTCCCTTGACATCAAGCTAAA 1084
OY 101 AenCyAspPheSerPheThrGlyLeuGlnHisValThrAspValIleIleMetLysLys 120
DB 1085 AATTGATTTTCTCTTTACTGACCTTCAACAGCTTACTGATTAATAAATAATATATAAAG 1144
OY 121 GluIleGlnGluGlyIleGlyLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 140
DB 1145 GAAAAAGAGAGAGATTGTAAGAGGGCAATCTGCTTCAGACAGACATTCCTGCC 1204
OY 141 ThrValGlnHisThrMetAlaCyHisIleuValLysArgThrHisArgAlaIleLeuPhe 160
DB 1205 ACGTACAGACACAAATGGCATGTCATCTGTGAAAAAACAATCGGGCATTCGTGTT 1264
OY 161 CysLysGlnArgAspLeuLeuProGlnAspAsnAlaValLeuValAlaSerGlyGlyVal 180
DB 1265 TGTAAACAGAGACATTTGTAACCTCAAAATATGACAGTACGTTGATCTGTGAGTGC 1324
OY 181 AlaSerAspPheThrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr 200
DB 1325 GCAAGTAACCTTCTGTATCCGACAGCTCTGAAATTTTAAACAAACGCAACAGATGCACT 1384
OY 201 LeuLeuCySerProProProArgLeuCyThrAspAsnGlyIleMetIleAlaTrpAsnGly 220
DB 1385 TTCTTGTGCTCTCTCCACACATATGACATGATATGCACTTAATGATTGCA----- 1435
OY 221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTrpGlu 240
DB 1435 ----- 1435
OY 241 ProLysCySerProLeuGlyValAspIleSerIleGlyValGlyValAlaSerIleLysVal 260
DB 1436 ---TGATGTCTCTTGTGAGTAAGACATATCAAAAGAGCTTGAGAGCTTCATATAAGTA 1492
OY 261 ProGlnLeuLysMetGluIle 267
DB 1493 CCAAAATTAATAATGAGATA 1513

RESULT 13
LOCUS BC078974 1546 bp mRNA linear ROD 03-AUG-2004
DEFINITION Rattus norvegicus cDNA clone IMAGE:7111906, partial cds.
ACCESSION BC078974
VERSION BC078974.1 GI:50926879
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (Norway rat)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 1546)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buettow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stephenson,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulv,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Heitton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,B.D.,

```

```

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,
Schmerch,A., Schein,J.B., Jones,S.J., and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1546)
Director MGC Project.
Direct Submission
Submitted (02-AUG-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Howard Jacobs
cDNA Library Preparation: Express Genomics
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNLT at: http://image.llnl.gov
Series: IRAX Plate: 182 Row: F Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
Location/Qualifiers
1. 1546
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:7111906"
/tissue_type="Testis, rat (Brown Norway)"
/clone_11b="NIH_MGC_237"
/lab_host="DH10B"
/notes="vector: pXpress1"
144. >1546
/codon_start=1
/product="Unknown (protein for IMAGE:7111906)"
/protein_id="AAH78974.1"
/db_xref="GI:50926880"
/translation="MLMSKTAGAIIPRPNSVNGFIRPNVOPRALFHKLVIGIET
SCDPTAAAVDETGNVGEALHSQTEVRLKGTGIVPVAQOLHENTORIVBEALSAS
GVSPSDLSAIVTTIRKGLASLVGGLSVOLVNOFKPFIPIHMEHMAITRLTLTK
VGFPFLVLLISGHCILALVOSVDFLLGLSLDAPDMLDKVARRSLIKHPECST
MSGKAIENLAKGNRFHTINPWNQKNCDFSTFGIQTHTDKLITKEKEBEGIKG
OILSSADIAAIVOHATACHTAKTRHAILRCQOKNLSPANAVLVGVASINLYIR
PALEIVANAQCTLLCPPRCTGNGIANGIERAAGILHDVDIRPAPAGS
IVSWLRVITATLIDSDVSSTHYVAHSPLNSGSRANIQTSMSCSTQIVMRTVHTLT
NINLKSXXXXXXXXXX"
ORIGIN
Alignment Scores:
Pred. No.: 4.7e-100 Length: 1546
Score: 1116.00 Matches: 214
Percent Similarity: 92.56% Conservative: 10
Best Local Similarity: 88.43% Mismatches: 18
Query Match: 80.58% Indels: 0
Gaps: 0
US-10-649-273-2_COPY_148_414 (1-267) x BC078974 (1-1546)
OY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
DB 585 ATGAGGCTCAGCGCTACATTTAGGCTGACCCCAAAAGTCGATTCCTTTTAGTT 644

```


QY 21 LeuLeuIIeserGIyGIySIcyleuLeuAlaleuValInGIyValIsarSpheLeu 40
 DB 645 CTTTTCATTTCTGAGAGCCATCTGCTGTTGGCGTTACTGACAGTGTTCAGATTTTCTG 704
 QY 41 LeuLeuGIlyserSerLeuaspIIeAlaProGIyAspMetLeuaspIysValAlaArg 60
 DB 705 CTCCTGGGAGATCCCTCGACATAGCCGACGACATGCTTGCAAGGTTGGCAAGAGA 764
 QY 61 LeuSerLeuIIelyshISProGIyCysSerThmetSerGIyGIyValaIleGIuHIS 80
 DB 765 CTTTCTTAATCAACATCCAGATGTTCTACATGATGCTGGGAAGCTATAGACAT 824
 QY 81 LeuAlaIySGInGIyAsnaArgPheHISpheaSpIIelysProProLeuHISHisAlaLys 100
 DB 825 TTGGCCAAAGAAAGAAATAGATTCACACTTACTATCAATCACCACCAAGAAATGTAAG 884
 QY 101 AsnCyAspPheSerPheThrGIyLeuGIuHISValThrAspIysIleIleMetLysLys 120
 DB 885 AACTGTGATTTTCTTTTACGGACCTTCACATGTCCACATTAAGCTAATACACACAG 944
 QY 121 GIuLysGIuGIuGIyIleGIuLysGIyGIuIleLeuSerSerAlaAlaaspIIeAla 140
 DB 945 GAAAGAAAGAAAGCACTTGAGAGGGCAATCTGTCAACGCCGACATTTGCTGCT 1004
 QY 141 ThrValGIuHISThrMetAlaCySHISLeuValIysArgThrHISArgAlaIleLeuPhe 160
 DB 1005 GCGGTACAGACGCCAACAGCGTGCCACTTGCGAAAGAAACATCGTGCTATTCGTTT 1064
 QY 161 CysLysGIuArgAspLeuLeuProGIuAsnAsnAlaValLeuValaIasergIyGIyVal 180
 DB 1065 TGCACACAGAAATATTTGCTATCTCCAGCTACACGACATATATGTTGTGTGGAGGTGT 1124
 QY 181 AlaSerAsnPheThrIleArgArgAlaLeuGIuIleLeuThrasnaIathrInGIyCysThr 200
 DB 1125 GCAAGTAACCTTGATCAATCCAGAGACATGGAATTTGTAGCAATGGAACCAATGCACT 1184
 QY 201 LeuLeuCySerProProArgLeuCySThrAspAsnGIyIleMetIleAlaIatrasnGIy 220
 DB 1185 TTGTTGTGTCCCTCCGACATGTCATGCAATGATGATGATGATGATGATGATGATGATGAT 1244
 QY 221 IIEGIuArgLeuArgAlaGIyLeuGIyIleLeuHISaspIIeGIuGIyIleArgTyrGIu 240
 DB 1245 ATGGAAGATTAAGTGTGCTGCTGGCATTTTACATGATGATGATGATGATGATGATGATGAT 1304
 QY 241 ProLys 242
 DB 1305 CCMAAG 1310
 RESULT 14
 BC051211 1109 bp mRNA linear ROD 15-APR-2003
 LOCUS Mus musculus, clone IMAGE:1327545, mRNA.
 DEFINITION BC051211
 ACCESSION BC051211 GI:29881634
 VERSION BC051211.1
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1109)
 Strusberg, R.
 Direct Submission
 Submitted (14-APR-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Marcello Bento Soares, Ph.D.
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMIL)
 DNA Sequencing by: Institute for Systems Biology

http://www.systemsbio.org
 contact: amadan@systemsbio.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Keteman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILMIL at: http://image.llnl.gov
 Series: IRAC Plate: 113 Row: D Column: 1
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.
 FEATURES
 source location/Qualifiers
 1..1109
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:1327545"
 /tissue_type="Thymus gland, mouse"
 /clone_id="Soares_thymus_2nbmt"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac"
 ORIGIN
 Alignment Scores:
 Pred. No.: 6,23e-84 Length: 1109
 Score: 950.50 Matches: 190
 Percent Similarity: 79.10% Conservative: 22
 Best Local Similarity: 70.90% Mismatches: 35
 Query Match: 68.63% Indels: 21
 DB: Gaps: 4
 US-10-649-273-2_COPY_148_414 (1-267) x BC051211 (1-1109)
 QY 18 PheLeuValLeuLeuIIeser-----GIyGIyHIScyleu----- 29
 DB 82 TTTTTCATTTCTGAGAGCCATCTGCTGTTGGCGTTACTGACAGTGTTCAGATTTTCTG 141
 QY 30 -----LeuAlaLeuValGIuGIyValIsarSpheLeuLeuGIyLysSerLeuasp 47
 DB 142 TCTGTTGAAACAGTCTGCTCAATGTTGGGTTATCTTGTCTTAAT-----AAT 192
 QY 48 IIEAlaProGIyAspMetLeuaspIys-----ValAlaArg 59
 DB 193 TTGCAGCCAAATATATTATTAAGAAAAAANGTAATGCTGTTTATATAGTGGCAAGA 252
 QY 60 ArgLeuSerLeuIIelyshISProGIyCysSerThmetSerGIyGIyValaIleGIu 79
 DB 253 AGACTTCTTAATCAACATCCAGATGTTCTCAATGATGATGATGATGATGATGATGATGATGAT 312
 QY 80 HisLeuAlaIySGInGIyAsnaArgPheHISpheaSpIIelysProProLeuHISHisAla 99
 DB 313 CAGTTGGCCAAAGAGCAATAGATTCATTTTCTACATCAACCTTACAGAAATCT 372
 QY 100 LysAsnCyAspPheSerPheThrGIyLeuGIuHISValThrAspIysIleIleMetLys 119
 DB 373 AAGAAATGCAATTTTCTTTCACGGACTTCAACATATATGATGATGATGATGATGATGATGAT 432
 QY 120 LysGIuLysGIuGIyIleGIuLysGIyGIuIleLeuSerSerAlaAlaaspIIeAla 139
 DB 433 AAGGAAAGAAAGAAAGCACTTGAGAGGGCAATTCGTATCATGCTGACGACATTTGCT 492
 QY 140 AlaThrValGIuHISThrMetAlaCySHISLeuValIysArgThrHISArgAlaIleLeu 159
 DB 493 GCTCGGTACAGCATGCAACGCTGCACTTGCGAAAGAAACACATCGGCTATTTGTG 552
 QY 160 PheCysLysGIuArgAspLeuLeuProGIuAsnAsnAlaValLeuValaIasergIyGIy 179
 DB 553 TTTTTCAGAGCAAGAAATTTGCTTCTTCACCTTAACGACATTAATGTTATTTGAGAGT 612
 QY 180 ValAlaSerAsnPheThrIleArgArgAlaLeuGIuIleLeuThrasnaIathrInGIy 199
 DB 613 GTTCAAGTAACCTTGATCAATCCGAAAGCAATTTGGAATTTGCGAAATGCAACGACGTGC 672

QY 200 ThrLeuLeuCyseProProProArgLeuCyseThraAspanglylleMetlleaATPasn 219
DB 673 AGCTGTGTGTCACCTCCAGACAGCTGTCAGTACAGATGATGATGATGATGAT 732
QY 220 GATTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 239
DB 733 GGAATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 792
QY 240 GATProLeuCyseProLeuGlyValAspLleSerLysGlyValGlyLysLleSerLys 259
DB 793 GAACCAAAAGTCTCTTGGATGATGATGATGATGATGATGATGATGATGATGAT 852
QY 260 ValProGlnLeuLysMetGluile 267
DB 853 GTACCCGATTAATAATGCACATT 876
RESULT 15
BX934991 1522 bp mRNA linear VRT 02-FEB-2004
LOCUS Gallus gallus finished cDNA, clone CHEST169114.
DEFINITION BX934991
ACCESSION BX934991 GI:41635519
VERSION BX934991.1 GI:41635519
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 1522)
Boardman, P. B., Bonfield, J. K., Brown, W. R. A., Carder, C., Chalk, S. E.,
Cronin, M. D. R., Davies, R. M., Francis, M. D., Grafham, D. V.,
Hubbard, S. J., Humphray, S. J., Hunt, P. J., Maddison, M., McLaren, S. R.,
Noble, D., Overton, J. M., Rogers, J., Scott, C. E., Taylor, R. G.,
Tickle, C. and Wilson, S. A.
Direct Submission
Submitted (02-FEB-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: chicken@hms.unist.ac.uk
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
sequencing project.
This sequence is from the
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
from a library constructed by Elizabeth Bosch. cDNA was prepared
from RNA extracted from muscle, normalised, and poly A-tailed.
ECORI-NotI cut cDNA was then ligated into the vector. Vector:
pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
coli DH10B.
FEATURES
source location/Qualifiers
1..1522
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="layer and broiler"
/db_xref="taxon:9031"
/clone="CHEST169114"
/clone_lib="CSEGRB11"
/dev_stage="adult"
ORIGIN
Alignment Scores:
Pred. No.: 1.24e-82 Length: 1522
Score: 939.00 Matches: 176
Percent Similarity: 81.06% Conservative: 38
Best Local Similarity: 66.67% Mismatches: 50
Query Match: 67.80% Indels: 0
DB: 5 Gaps: 0
US-10-649-273-2_COPY_148_414 (1-267) x BX934991 (1-1522)
QY 1 MetGluAlaHisAlaLeuThrLleArgLeuThrAsnLysValGluPheProPheLeuVal 20
DB 587 ATGAGAGCTCAGCAGCTTACATCAGACTGACAGAGCAGAGATTTCCCTTCTAGTT 646
QY 21 LeuLeuLleSerGlyGlyHisCyseLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40

DB 647 CTTTACTCTCGAGAGCTGACATCTTGGACAGTACAGACAGAGATTTCAGATTCTT 706
QY 41 LeuLeuGlyLysSerLeuAspLleAlaProGlyAspMetLeuAspLysValAlaArg 60
DB 707 CTGCTGGACAGCTCATGATATAGCAGACAGTGCATGCTGGATTAAGATGACAGAG 766
QY 61 LeuSerLeuLleLysHisProGlnCyseSerThrMetSerGlyGlyLysAlaLleGlyHis 80
DB 767 CTTCTTTAGTGAAGACACCCGAGAGCCAGAGCATGCTGGGGGAGAGCATTAAGACAC 826
QY 81 LeuAlaLysGlnGlyAsnArgPheHisAspLleLysProProLeuHisAlaLys 100
DB 827 CTGGCTCAACCGAGAGCTGCAACAGTACCTTCAACCTTCCATGCAACAGTATGCT 886
QY 101 AsnCyseAspPheSerPheThrGlyLeuGlnHisValThrAspLysLlelleMetLysLys 120
DB 887 AACTGTGATTTTCTTTCTCCGACTTACAGACCTTTCACAAAGCCATTCTTCAGAA 946
QY 121 GluLysGlnGlyLysLleGlyLysGlyGlnLleLeuSerSerAlaAlaAspLleAla 140
DB 947 GAAAAGAGAGAGATTCAGAGAGGGAATCTGCTGCTGCTTAAGACATCTGCTGCT 1006
QY 141 ThrValGlnHisThrMetAlaCyseHisLeuValLysArgThrHisArgAlaLleLeuPhe 160
DB 1007 GCTGCACAGCAGTATGCTGCTGCTCATATATCCAGCCGACACACAGCCATGCTCTTC 1066
QY 161 CysLysGlnArgAspLeuLeuProGlnHisAsnAlaValLleValAlaSerGlyGlyVal 180
DB 1067 TGCATGAAAACAGCATATTTATACCAAAAACCTGCACCTGTGTGTATCAGAGAGATT 1126
QY 181 AlaSerAsnPheThrLleArgArgAlaLeuGlnLleLeuThrAsnAlaThrGlnCysThr 200
DB 1127 GCAAGTATTCAGTATATCAGAAAAGAGCTGCAGACTGCGCAAAAGCGTTTGGCT 1186
QY 201 LeuLeuCyseProProProArgLeuCyseThraAspanglylleMetlleaATPasn 220
DB 1187 TTTCTGTCTCTCCCTCCAGAGCTGCGACCGATATGTTATGATTCATGAGATGCTC 1246
QY 221 lleGluArgLeuArgAlaGlyLeuGlyLleLeuHisAspLlelleGlyLysLleArgTyrg 240
DB 1247 ATGAAAGGTGCGCGAGAGTGTGTTATATCAGTACTGAAGCGATTCGCTACGAA 1306
QY 241 ProLysCyseProLeuGlyValAspLleSerLysGlyValGlyLysLleSerLysVal 260
DB 1307 CCAAAAGCTCCCTTGAATGATATTTCCAAAAGAGTTGAAGAGATTCCATCAAGTG 1366
QY 261 ProGlnLeuLys 264
DB 1367 CCAAGACTTAAG 1378

Search completed: June 16, 2005, 23:18:39
Job time : 3202.26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 16, 2005, 12:03:59 ; Search time 385.989 Seconds
(without alignments)
4094.859 Million cell updates/sec

Title: US-10-649-273-2_COPY_148_414

Perfect score: 1385

Sequence: 1 MEAHALTRITNKVEFFLV.....DISKEVGASIKVPOLKMEI 267

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Delop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DRV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10649273/runat_15062005_111415_6022/spp_query.fasta_1.1429
-DB=N Geneseq_16Dec04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10649273.@CGN_1.1.1063.@runat_15062005_111415_6022 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004as:*\n13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1385	100.0	1416	8	ABX70950
2	1385	100.0	1526	6	ABST76639
3	1385	100.0	2058	6	ABA93268
4	1385	100.0	2197	6	ABST76635
5	1385	100.0	2572	8	ABT23207

6	1348	97.3	1820	6	AAD46856	Aad46856 Human gly
7	1348	97.3	1820	7	ACA60887	ACA60887 Human CDN
8	1348	97.3	1821	10	ABSS7020	ABSS7020 cDNA enco
9	1204	86.9	2208	10	ADA52832	ADA52832 Human cod
10	1204	86.9	2890	12	ADQ24627	ADQ24627 Human sof
11	995.5	71.9	3358	10	ADE31345	Ade31345 Human dia
12	983	71.0	1572	6	ABQ75508	Abq75508 Murine si
13	599	43.2	2734	6	AAS84622	Aas84622 DNA enco
14	468	33.8	371	12	ADL66726	Adl66726 DNA up-re
15	468	33.8	371	12	ADL66726	Adl66726 DNA up-re
16	415.5	30.0	1601	4	ABL24632	Abi24632 Drosoephil
17	415.5	30.0	1656	3	AAC38454	Aac38454 Arabidops
18	401.5	29.0	1557	3	AAD48239	Aad48239 Ehrlichia
19	342	24.7	4360	6	AAD48239	Aad48239 Human CDN
20	338.5	24.4	1385	4	AAH15110	Aah15110 Human gen
21	321.5	23.2	1146	8	ACA26804	ACA26804 Prokaryot
22	314.5	22.7	936	8	ACA20445	ACA20445 Prokaryot
23	308.5	22.3	1044	8	ACA39102	ACA39102 Prokaryot
24	308.5	22.2	94750	4	AAE28551	Aae28551 Genomic f
25	307	22.2	1053	12	ADL03120	Adl03120 DNA enco
26	304.5	22.0	1032	8	AAE91424	Aae91424 Moraxella
27	301.5	21.8	1000	4	AAE91424	Aae91424 Moraxella
28	301.5	21.8	1000	6	ABK37804	Abk37804 DNA sequ
29	300	21.7	1092	6	ABQ90383	Abq90383 M. capul
30	291.5	21.0	1026	4	AAS54064	Aas54064 Pseudomon
31	291.5	21.0	1026	8	ACA42146	ACA42146 Prokaryot
32	291.5	21.0	1026	10	ADG73141	Adg73141 P aerugin
33	291.5	21.0	1059	11	ABD02280	Abd02280 Pseudomon
34	291.5	21.0	1206	11	ABD02197	Abd02197 Pseudomon
35	289.5	20.9	1026	10	ADG73141	Adg73141 P aerugin
36	289.5	20.9	1029	8	ACA34150	Aca34150 Haemophil
37	289.5	20.9	1029	4	AAS53309	Aas53309 Haemophil
38	289.5	20.9	110000	2	AAT42063_05	Continuaction (6 of
39	285	20.6	372	5	ABV61069	Abv61069 Human pro
40	282.5	20.4	9967	13	ADT05493	Adt05493 Haemophil
41	282.5	20.4	85814	13	ADT05644	Adt05644 Haemophil
42	280.5	20.3	1020	8	ACA43484	Aca43484 Prokaryot
43	280.5	20.3	1074	10	ADP02056	Adp02056 Bacterial
44	279.5	20.2	1014	4	AAS56045	Aas56045 Salmonell
45	279.5	20.2	1014	8	ACA51431	Aca51431 Prokaryot

ALIGNMENTS

RESULT 1	ABX70950	standard; cDNA; 1416 BP.
ID	ABX70950	standard; cDNA; 1416 BP.
AC	ABX70950;	
XX		
DT	05-MAR-2003	(first entry)
XX		
DE	Novel human cDNA sequence #175.	
XX		
KW	Human; gene; ss; nervous system disorder; peripheral neuropathy;	
KW	Huntington's disease; amyotrophic lateral sclerosis; haemophilia;	
KW	neurodegenerative disease; Parkinson's disease; Alzheimer's disease;	
KW	autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;	
KW	insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;	
KW	ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;	
KW	fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;	
KW	coagulation disorder; cancer; tumour; inflammatory disease; septic shock;	
KW	Crohn's disease; anaplasia; proliferation; chemotactic;	
KW	differentiation; stem cell growth factor; hematopoiesis; chemokine;	
KW	haemostatic; antiinflammatory; expressed sequence tag; ESR.	
OS	Homo sapiens.	
XX		
PN	W0200281731-22.	
XX		
PD	17-OCT-2002.	
XX		
PF	29-JAN-2002; 2002WO-US001222.	

XX 30-JAN-2001; 2001US-00774528.
 PR (HYSE-) HYSEO INC.
 PA (GOOD/) GOODRICH R. W.
 PA
 XX Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;
 XX MPI; 2003-058563/05.

XX Novel polypeptide useful for treating neurodegenerative diseases, myeloid
 PT or lymphoid cell disorders, bone disorders, mechanical and traumatic
 PT disorders, coagulation disorders, and inflammatory diseases.

XX Claim 1; Page: 612pp; English.

XX This invention relates to the cDNA sequences encoding an isolated novel
 CC human polypeptide. The protein encoded by the nucleic acid of the
 CC invention is useful for treating central and peripheral nervous system
 CC diseases (e.g. peripheral neuropathy, Huntington's disease, amyotrophic
 CC lateral sclerosis); neurodegenerative diseases (e.g. Parkinson's disease,
 CC Alzheimer's disease); autoimmune disease (e.g. systemic lupus
 CC erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus)
 CC ; myeloid or lymphoid cell disorders (e.g. anaemia and thrombocytopaenia)
 CC ; wounds, ulcers, burns; bone disorders (e.g. osteoporosis,
 CC osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head
 CC trauma); lung or liver fibrosis; reperfusion injury in various tissues;
 CC bacterial, viral or fungal infections; allergic conditions such as
 CC allergic rhinitis, asthma; coagulation disorders (e.g. haemophilia);
 CC cancer and tumors; and inflammatory diseases (e.g. septic shock, Crohn's
 CC disease, anaphylaxis). The protein may be used to inhibit the growth,
 CC infection or function of infectious agents such as bacteria, fungi,
 CC viruses, or to effect bodily characteristics, biorhythms or circadian
 CC cycles of rhythms. The protein may also have
 CC proliferation/differentiation, stem cell growth factor, haematopoiesis
 CC regulation, immune stimulation or suppressing, chemotactic/chemokinetic,
 CC haemostatic and thrombolytic, receptor/ligand, and antiinflammatory
 CC activities. The cDNA sequences of the invention are useful for expressing
 CC recombinant protein for analysis. The present sequence represents a novel
 CC human cDNA sequence of the invention, this sequence is an expressed
 CC sequence tag (EST) and was identified using subtractive hybridization

XX Sequence 1416 BP; 441 A; 280 C; 273 G; 422 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.: 4,68e-147 Length: 1416
 Score: 1385.00 Matches: 267
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x ABX70950 (1-1416)

QY 1 MetGUAAlAHISAlAeuthrileagLeuthRasnlYsValGluPheProPheLeuVal 20
 DB 502 ATGGAGGCTCATGCACTTACTATTAGGTTACCAATTAAGTAGAATTTCTTTTGT 561
 QY 21 LeuLeuIISeSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValISeSerPheLeu 40
 DB 562 CTTTTCATTTCTGGAGGTCACCTGCTGTTGGCATTAAGGAGTTTCAGATTTCGTG 621
 QY 41 LeuLeuGlyIySeSerLeuAspIleAlaProGlyAspMetLeuAspIyValAlaArg 60
 DB 622 CTTCTTGGAAAGCTTTGGACATAGACACAGGTGACATGCTGACCAAGTGGCAAGA 681
 QY 61 LeuSerLeuIIleYshIspProGluCySeSerThwecSerGlyGlyValAlaIleGluHis 80
 DB 682 CTTTCTTTATATAAACATCCAGAGTGTCTCCACATGAGTGCGGGAAGCATGAAAT 741
 QY 81 LeuAlaIyGlnGlyAsnArgPheHisPheAspIleYsPheProPheLeuHisAlaIyS 100

DB 742 TTGGCCAAACAGGAATAGATTTCATTGGACATCAAACTCCCTTCATCATGCTTAA 801
 QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspIyIleIleMetIyS 120
 DB 802 AATTGTGATTTTCTTTTACTGACCTTCAACACTTCTGATTAATAATATGAAAAAG 861
 QY 121 GluIySGluGluGlyIleGluIySGlyGlnIleuSerSerAlaAlaAspIleAla 140
 DB 862 GAAAAGAGGAGATATGGAAAGGGCAATCTGCTTCAGACAGACATGCTGCC 921
 QY 141 ThrValGlnHisThrMetAlaCysHisLeuValIySerGlyThrHisAlaGlnIleuPhe 160
 DB 922 ACAGTACAGCACACAAAGGACATGCTGTGGAAGAACACATCGGCTATTCGTTT 981
 QY 161 CysIySGlnArgAspIleuLeuProGlnAsnAsnAlaValLeuValAlaSeSerGlyVal 180
 DB 982 TGTAAAGAGAGAGCTGTGTAACCTCAAAATATGACATGCTGTTGCATCTGTGTTTC 1041
 QY 181 AlaSerAspPheIyIleArgArgAlaLeuGlnIleuThrAsnAlaThrGlnCysThr 200
 DB 1042 GCAAGTAACTTCTATATCCGAGAGCTGTGGAATTTTAAACAAAGCAACAGTCACT 1101
 QY 201 LeuLeuCySPProPProArgLeuCyThrAspAsnGlyIleMetIleAlaTPAsnGly 220
 DB 1102 TTGTTGTCTCTCTCCAGACTATGACATGATATATGATATGATGATGATGATGAT 1161
 QY 221 ILeuGlnArgLeuArgAlaGlyLeuGlyIleuHisAspIleGluGlyIleArgIyGlu 240
 DB 1162 ATTAAAGACTACCTGCTGCTGGCAATTTTACATGACATAGAAAGCATCCGCTATATA 1221
 QY 241 ProIySCySPLeuGlyValAspIleSerIySGlyValGlyGlnAlaSerIleYsVal 260
 DB 1222 CCAAAATCTCTCTTGGAGTAGACATATCAAAAGATGGAGAGCTTCCATAAAGTA 1281
 QY 261 ProGlnLeuIySMetGluIle 267
 DB 1282 CCACAAATTAATAAATGAGATTA 1302
 DB
 XX RESULT 2
 XX ABS76639
 XX ID ABS76639 standard; DNA; 1526 BP.
 XX AC ABS76639;
 XX
 XX 11-DEC-2002 (first entry)
 XX
 XX DNA encoding novel human metalloprotease MPI fragment #1.
 XX Metalloprotease; MP-1; immune disorder; glutamate transport; cancer;
 XX motor neuron disorder; amyotrophic lateral sclerosis; AIDS; diabetes;
 XX reproductive disorder; Klinefelter's syndrome; germinal cell aplasia;
 XX genital wart; metabolic disorder; premature puberty; Kallman syndrome;
 XX Cushing's syndrome; neurodegenerative disease; Alzheimer's disease;
 XX Parkinson's disease; Huntington's disease; Tourette syndrome; sepsis;
 XX liver disease; renal disease; immune disorder; rheumatoid arthritis;
 XX acquired immunodeficiency syndrome; AIDS; pulmonary disease; pneumonia;
 XX emphysema; cystic fibrosis; vascular disorder; inflammatory disorder;
 XX neurological disorder; gene; ds.
 XX Homo sapiens.
 XX W0200272751-A2.
 XX
 XX 19-SEP-2002.
 XX
 XX 05-FEB-2002; 2002WO-US003353.
 XX
 XX 05-FEB-2001; 2001US-0266518P.
 XX 10-APR-2001; 2001US-0282814P.
 XX
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 XX Chen J, Feder J, Nelson TC, Duclos F, Kryetsek S;

XX WPI; 2002-723329/78.
 DR P-PSDB; ABG96487.
 XX
 PT New isolated nucleic acid encoding MP-1 protein, useful for preventing,
 PT treating, or ameliorating diseases associated with aberrant
 PT metalloproteinase activity, e.g. immune, metabolic, inflammatory and
 PT neurological disorders.
 XX
 PS Disclosure; Page 462-463; 473pp; English.
 XX
 CC The invention describes an isolated nucleic acid molecule (1) encoding a
 CC metalloproteinase (MP-1). (1) is useful for preventing, treating, or
 CC ameliorating a medical condition, particularly an immune disorder, an
 CC aberrant glutamate transport or motor neuron disorder, such as
 CC amyotrophic lateral sclerosis (ALS), its juvenile form or an ALS-like
 CC condition. The compositions and methods are also useful for diagnosing,
 CC prognosticating, treating, ameliorating and/or treating disorders
 CC associated with MP-1 activity, e.g. diabetes, cancer, reproductive
 CC disorders (e.g. Klinefelter's syndrome, genital warts, or germinal cell
 CC aplasia), metabolic disorders (e.g. premature puberty, Kallman syndrome,
 CC or Cushing's syndrome), neurodegenerative diseases (Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease or Tourette syndrome), liver
 CC and renal diseases and immune disorders (e.g. AIDS, Rheumatoid arthritis
 CC or sepsis), pulmonary diseases (e.g. pneumonia, emphysema or cystic
 CC fibrosis) and vascular, inflammatory and neurological disorders (e.g.
 CC Alzheimer's disease or Parkinson's disease). This sequence represents a
 CC metalloproteinase MPI polynucleotide
 XX
 SQ Sequence 1526 BP; 484 A; 297 C; 289 G; 456 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,2e-147 Length: 1526
 Score: 1385.00 Matches: 267
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x ABS76639 (1-1526)

QY 1 MetGluAlaHisAlaLeuThrFlaArgLeuThrAsnValGluPheProPheLeuVal 20
 DB 1 ATGGAGGCTCATGCACTTACTTATAGTTGACCAATAAGAGCAATTTCTTTTATGTT 60
 QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
 DB 61 CTTTGTATTTCTGGAGGTCACTGTCGTGGCATTAGTTCAAGAGTTTCGATTTCTCG 120
 QY 41 LeuLeuGlyIleSerLeuAspPheLeuAlaProGlyIleAspPheLeuValAlaArg 60
 DB 121 CTTCTTGAAAGCTTTGGACATAGCACACAGTGCATGCTTGACCAAGTGGCAAGAA 180
 QY 61 LeuSerLeuIleValSerGlyIleSerGlyIleValSerGlyIleValSerGlyIleVal 80
 DB 181 CTTTCTTAATAAACAATCCAGAGTCTCCACATGATGGTGGGAAAGCCATGACAT 240
 QY 81 LeuAlaIleGlnGlyAsnArgPheHisAspPheLeuValProLeuHisHisAlaVal 100
 DB 241 TTGGCCCAACAGAAATAGATTTCATTTGACATCAACCTCCCTTCATCATGCTAA 300
 QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspValIleIleMetLeuVal 120
 DB 301 AATTGTATTTCTTTTACTGACCTTCAACACCTTACTGATTAATAATATATGAAAAAG 360
 QY 121 GlnLeuGlnGlnGlyValIleGlnGlyGlnIleLeuSerSerAlaAlaAspIleAla 140
 DB 361 GAAAGAAAGAGAGGATATGAGAGGGGCAATCCGCTTGACGACGACATGCTCTCC 420
 QY 141 ThrValGlnHisThrMetAlaCysHisLeuValIleArgThrHisArgAlaIleLeuPhe 160
 DB 421 ACGTACAGCACACATGACATGATCTTTGTGAAAAAGAACATCGGGCTATTTCTGTTT 480

QY 161 CysIleGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyVal 180
 DB 481 TGTAAAGCAGAGACTTGTACTTCAATAATATGACATGCTGTGCATGTGGTGTGC 540
 QY 181 AlaSerAsnPheValIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr 200
 DB 541 GCAAGTAACTTCTATATCCGACAGCTTGAAATTTTAAACAGCAACAGAGTCACT 600
 QY 201 LeuLeuCysProProProArgLeuCysThrAspAsnGlyIleMetIleAlaTrpAsnGly 220
 DB 601 TTGTTGTCTCTCCCTCCAGACTATGCACTGATATGACATTATGATTCATGAAATGCT 660
 QY 221 IleGlnArgLeuAlaGlnGlyLeuGlnIleLeuHisAspIleGlnGlyIleArgGlyVal 240
 DB 661 ATTTGAAGACATACGCTGCTGGCATTTTAACTATGACATAGAAAGGACATCGCTATGAA 720
 QY 241 ProIleCysProLeuGlnValIleAspIleSerGlyValGlnAlaSerIleLeuVal 260
 DB 721 CCAAAATCTCTCTTGGAGTGAACATATCAAAAGAGTTGGAGAACTTCCATTAAGTA 780
 QY 261 ProGlnLeuIleValMetGluIle 267
 DB 781 CCACATTTAAATAATGAGATTA 801
 QY
 DB
 XX
 AC ABA93268; standard; cDNA; 2058 BP.
 XX
 AC ABA93268;
 XX
 DT 19-APR-2002 (first entry)
 XX
 DE Human O-sialoglycoproteinase-like protein encoding cDNA SEQ ID NO:1.
 XX
 XX Human, O-sialoglycoproteinase-like protein; OSGPLP; enzyme; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 110..1354
 FT /*tag= a
 FT /product= "O-sialoglycoproteinase-like protein"
 XX
 PN CN1318550-A.
 XX
 PD 24-OCT-2001.
 XX
 PF 19-APR-2000; 2000CN-00106834.
 XX
 PR 19-APR-2000; 2000CN-00106834.
 XX
 PA (SHAN-) SHANGHAI SHENGYUAN GENE DEV CO LTD.
 XX
 PI Mao Y, Xie Y;
 XX
 DR WPI: 2002-115090/16.
 XX
 DR P-PSDB; ABB05481.
 XX
 PT O-sialoglycoproteinase-like protein and encoding polynucleotide, useful
 PT for diagnosing, preventing and treating related diseases.
 XX
 PS Claim 5; Page 29-30 (Disclosure); 38pp; Chinese.
 XX
 CC The present sequence encodes human O-sialoglycoproteinase-like protein
 CC (OSGPLP). The present invention also describes: (1) the preparation of
 CC the OSGPLP protein; (2) applying the OSGPLP protein in diagnosis; (3) the
 CC prevention and/or treatment of related diseases; (4) utilizing the OSGPLP
 CC protein in screening its agonist, excitomotor and inhibitor and preparing
 CC an antibody against the OSGPLP protein; and (5) the use of the OSGPLP
 CC polynucleotide sequences, proteins, agonists, excitomotors, inhibitors
 CC and antibodies in treating diseases related to the abnormal OSGPLP gene
 CC and in preparing the medicine composite for the treatment

Sequence 2058 BP; 637 A; 400 C; 410 G; 611 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7,97e-147	Length:	2058
Score:	1385.00	Matches:	267
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-649-273-2_COPY_148_414 (1-267) x ABA93268 (1-2058)

```
OY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
DB 551 ATGGAGGCTCATGCACTTACTATTAGGTGACCAATTAAGTAGAATTTCTTTTAAATT 610
OY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
DB 611 CTTTGAATTTCTGAGAGTCACTGCTGTTGGCAATTACATGAGAGATTTCAGATTTTCTG 670
OY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 60
DB 671 CTTCTTGAAGAAGCTTTGACATAGCACAGGTGACATGCTTGAACAGGTGGCAAGAGA 720
OY 61 LeuSerLeuIleLysHisProGlyCysSerThrMetSerGlyGlyLysAlaIleGluHis 80
DB 731 CTTTCTTAATAAATCAATCCAGAGTCTCCACCATGATGCTGGGAAGCCATAGAACAT 790
OY 81 LeuAlaLysGlnGlyAsnAlaGlyPheHisPheAspIleLysProPheLeuHisAlaLys 100
DB 791 TTGGCCAAACAAGAAATAGATTTCATTTGACATCAAACTCCCTTGACATGCTAA 850
OY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
DB 851 AATTGATTTTCTTTTACTGACCTTCAACAGCTTACTATTAATAAATAAATGAAG 910
OY 121 GluLysGlnGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaAla 140
DB 911 GAAAAAGAGAGATTTGAGAAAGGGCAATCTGCTTCAGACAGACATTTGCTGCC 970
OY 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
DB 971 ACAGGTACACACACATATGCGATGCTCTGTGAAAGAACACATCGGCTATTCTGTT 1030
OY 161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyVal 180
DB 1031 TGTAAACAGAGACTGTTACTCTCAAAATATGCACTACTGCTGATCGTGGTGC 1090
OY 181 AlaSerAsnPheTyrlleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr 200
DB 1091 GCAAGTAACTTCTATATCCGACAGCTCTGAAATTTTAAACAAAGCAACAGTGCAT 1150
OY 201 LeuLeuCysProProProArgLeuCysThrAspAsnGlyIleMetIleAlaIleArgAsnGly 220
DB 1151 TTGTTGTGTCCTCTCCACAGACTATGACATGATATGCACTTATGATGATGAAGGT 1210
OY 221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyGln 240
DB 1211 ATTGAAAGACTAGCTGCTGCGCATTTTACATACATAGACAGCATCCGCTATGAA 1270
OY 241 ProLysCysProLeuGlyValAspIleSerLysGlnValGlyGlnAlaSerIleLysVal 260
DB 1271 CCAAAATGTCCTCTTGAGATGACATATCAAAAGAAAGTGGAGAGCTTCATTAAGTA 1330
OY 261 ProGlnLeuLysMetGluIle 267
DB 1331 CCACATTAATAATGAGATTA 1351
```

RESULT 4
ABS76635
ID ABS76635 standard; DNA; 2197 BP.
XX
AC ABS76635;

XX 11-DEC-2002 (first entry)
DT
XX DNA encoding novel human metalloprotease MP1.

XX Metalloprotease; MP-1; immune disorder; glutamate transport; cancer;
DE motor neuron disorder; amyotrophic lateral sclerosis; ALS; diabetes;
XX reproductive disorder; Klinefelter's syndrome; germinal cell aplasia;
KW genital wart; metabolic disorder; premature puberty; Kallman syndrome;
KW Cushing's syndrome; neurodegenerative disease; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; Tourette syndrome; sepsis;
KW liver disease; renal disease; immune disorder; rheumatoid arthritis;
KW acquired immunodeficiency syndrome; AIDS; pulmonary disease; pneumonia;
KW emphysema; cystic fibrosis; vascular disorder; inflammatory disorder;
KW neurological disorder; gene; ds.

XX Homo sapiens.

XX W0200272751-A2.

XX 19-SEP-2002.

XX 05-FEB-2002; 2002MO-US003353.

XX 05-FEB-2001; 2001US-0266518P.

XX 10-APR-2001; 2001US-0282814P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Chen J, Feder J, Nelson TC, Duclos F, Krystek S;

XX WPI: 2002-723329/78.

XX P-PSDB; ABG96478.

XX New isolated nucleic acid encoding MP-1 protein, useful for preventing,
PT treating, or ameliorating diseases associated with aberrant
PT metalloproteinase activity, e.g. immune, metabolic, inflammatory and
PT neurological disorders.

XX Claim 1; Fig 1A-C; 473pp; English.

XX The invention describes an isolated nucleic acid molecule (i) encoding a
CC metalloprotease (MP-1) (i) is useful for preventing, treating, or
CC ameliorating a medical condition, particularly an immune disorder, an
CC aberrant glutamate transport or motor neuron disorder, such as
CC amyotrophic lateral sclerosis (ALS), its juvenile form or an ALS-like
CC condition. The compositions and methods are also useful for diagnosing,
CC prognosticating, treating, ameliorating and/or treating disorders
CC associated with MP-1 activity, e.g. diabetes, cancer, reproductive
CC disorders (e.g. Klinefelter's syndrome, genital warts, or germinal cell
CC aplasia), metabolic disorders (e.g. premature puberty, Kallman syndrome,
CC or Cushing's syndrome), neurodegenerative diseases (Alzheimer's disease,
CC Parkinson's disease, Huntington's disease or Tourette syndrome), liver
CC and renal diseases and immune disorders (e.g. AIDS, rheumatoid arthritis
CC or sepsis), pulmonary diseases (e.g. pneumonia, emphysema or cystic
CC fibrosis) and vascular, inflammatory and neurological disorders (e.g.
CC Alzheimer's disease or Parkinson's disease). This sequence represents a
CC metalloprotease MP1 polynucleotide

XX Sequence 2197 BP; 681 A; 441 C; 439 G; 636 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8,75e-147	Length:	2197
Score:	1385.00	Matches:	267
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-649-273-2_COPY_148_414 (1-267) x ABS76635 (1-2197)

```
OY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
  |||
```


Db 672 ATGGAGGCTTCAGCTACTTACTATTAGGTGACCAATAAGTAGATTCTTTTAGTT 731
 QY 21 LeuLeu11SerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
 Db 732 CTTTATATTTCTGAGAGCTCTGCTCTTGGCAATTAAGAGATTGAGATTCTG 791
 QY 41 LeuLeuGlyLysSerLeuAsp11Lea1aProGlyAspPheLeuAspLysValAlaArgArg 60
 Db 792 CTTCTTGAAAGCTTTGACATAGCACAGGTGACATGCTTGACAAAGTGGCAAGAAGA 851
 QY 61 LeuSerLeu11LeuLysHisProGluGlyCysSerThrMetSerGlyGlyValAla11GlnHis 80
 Db 852 CTTTCTTTATATAAACAATCCAGAGTGTCTCCACCACTGAGTGTGGAAAGCCATGAAACAT 911
 QY 81 LeuAlaLysGlnGlyAsnArgPheHisPheAsp11LeuLysProPheLeuHisAlaLys 100
 Db 912 TTGGCCAAACAAAGAAATATATATTCATTGTGACATCAAACTCTCCCTGACATGCTTAA 971
 QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLys11LeuMetLysLys 120
 Db 972 AATTGATTTTTCTTTTACTGACCTTCAACAGCTTCTGATTAATAATATATGAAAAAG 1031
 QY 121 GlnLysGlnGluGly11LeuGlyGlyGln11LeuSerSerAlaAlaAsp11Lea1aAla 140
 Db 1032 GAAAGAGAGAGATATTGAGAAAGGAGCAATCTGTCTTCAGACAGACATTTGCTGCC 1091
 QY 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisSerGlyAla11LeuPhe 160
 Db 1092 ACGGTACAGACACAAATGAGATGATCTGTGTAAAGAAACACATCGGCTATTCCTGTT 1151
 QY 161 CysLysGlnArgAspLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyVal 180
 Db 1152 TGTAAGCAGAGAGACTTTTACTCTCAAAATATATCACTACGTGCTGCTGCTGCTGCTC 1211
 QY 181 AlaSerAsnPheTyrl1eArgArgAlaLeuGln11LeuThrAsnAlaThrGlnCysThr 200
 Db 1212 GCAAGTACTCTTATATATCCGACAGCTCTGGAATTTTAAACAAACGACACAGTCACT 1271
 QY 201 LeuLeuCysProProPheArgLeuGlyThrAspAsnGly11Met11Leu11TrpAsnGly 220
 Db 1272 TTGTTGTCTCTCTCCACAGCTATGACATGATATGATGATGATGATGATGATGATGAT 1331
 QY 221 11GlnArgLeuArgAlaGlyLeuGly11LeuHisAsp11LeuGly11ArgTyrGln 240
 Db 1332 ATTGAAGACTACGCTGCTGGGCAATTTTACATGATCATTAGAAAGGATCCGCTGAA 1391
 QY 241 ProLysCysProLeuGlyValAlaSer11SerLysGlnValGlyAlaSer11LeuVal 260
 Db 1392 CCAAAATGCTCTCTTGAGATGACATATCAAAAGAAAGTTGAGAGAACTTCCATATAAAGTA 1451
 QY 261 ProGlnLeuLysMetGluLeu 267
 Db 1452 CCACAATTAATAAATGAGAGATA 1472
 RESULT 5
 ABT23207
 ID ABT23207 standard; DNA; 2572 BP.
 XX
 XX ABT23207;
 DT 01-MAY-2003 (first entry)
 DE Human protein modification + maintenance molecule DNA SEQ ID No 36.
 XX
 KW Cytoskeletal; antiarteriosclerotic; anti-HIV; antiallergic; anticonvulsant;
 KW cerebroprotective; antiparkinsonian; nootropic; antiinflammatory;
 KW antitumor; hepatotropic; gynecological; antibacterial; virucide;
 KW protozoacide; antiparasitic; cell proliferative disease; PMOD;
 KW protein modification and maintenance molecule; immunogenic fragment;
 KW cancer; autoimmune; inflammatory disease; neurological disorder;
 KW gastrointestinal; developmental; vesicle trafficking disorder; infection;
 KW protein-protein interaction; drug-target interaction;
 KW gene expression profile; human; gene; ds.

XX Homo sapiens.
 OS
 XX MO2003000844-A2.
 PN
 XX 03-JAN-2003.
 PD
 XX 18-JUN-2002; 2002MO-US019360.
 PF
 XX 22-JUN-2001; 2001US-0300508P.
 PR 06-JUL-2001; 2001US-0303445P.
 PR 13-JUL-2001; 2001US-0305405P.
 PR 09-AUG-2001; 2001US-0311442P.
 PR 24-AUG-2001; 2001US-0314821P.
 PR 29-AUG-2001; 2001US-0315922P.
 PR 03-MAY-2002; 2002US-0378205P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Gandhi AR, Kable AE, Swarnakar A, Hafalia AUA, Tran B, Duggan BM,
 PI Warren BA, Ison CH, Honchell CD, Nguyen DB, Lu DM, Lee EA, Yue H,
 PI Forsythe TJ, Barroso I, Rankumar J, Griffin JA, Li JK, Yang J;
 PI Thangavelu K, Gietzen KJ, Ding L, Baughn MR, Borowsky ML, Yao MG,
 PI Walla NK, Mason PM, Gururajan R, Lee S, Becha SD, Lee ST, Tran UK,
 PI Elliott VE, Luo W, Sprague WW, Tang YT, Lu Y, Zebartjadian Y;
 XX WPI; 2003-184039/18.
 DR P-PSDB; ABJ26654.
 XX
 PT New isolated human PMOD polypeptide and polynucleotide, useful for
 PT diagnosing, treating and preventing diseases or conditions associated
 PT with the aberrant PMOD expression, e.g. cancer, AIDS, atherosclerosis and
 PT infections.
 PT
 XX
 PS Claim 91; Page 211; 225pp; English.
 PS
 XX The invention relates to an isolated polypeptide comprising: any of 28
 CC sequences of 48-1256 amino acids; a natural amino acid sequence at least
 CC 90% identical to the 28 amino acid sequences, 94% identical to a sequence
 CC of 703 or 267 amino acids, 96% identical to a sequence of 414 amino
 CC acids, or 97% identical to a sequence of 242 amino acids, all given in
 CC the specification; or a biologically active or immunogenic fragment of
 CC the isolated polypeptide. The polypeptides and polynucleotides are useful
 CC in diagnosing, treating and preventing diseases or conditions associated
 CC with the decreased expression of protein modification and maintenance
 CC molecules (PMOD), such as cell proliferative diseases (e.g. cancer,
 CC atherosclerosis), autoimmune/inflammatory diseases (e.g. AIDS,
 CC allergies), neurological disorders (e.g. stroke, Parkinson's disease,
 CC epilepsy), gastrointestinal (e.g. ulcer, cirrhosis), reproductive (e.g.
 CC endometriosis), developmental, vesicle trafficking disorders, and
 CC infections (e.g. bacterial, viral, parasitic, protozoal). These are also
 CC useful in assessing the effects of exogenous compounds on the expression
 CC of nucleic acid and amino acid sequences of PMOD. The PMOD or its
 CC fragments are useful in screening compounds for effectiveness as agonist
 CC or antagonist of the polypeptides, or in altering the expression of the
 CC target polynucleotide and compounds that specifically bind to or modulate
 CC the activity of the polypeptide. The microarray is useful in monitoring
 CC or measuring protein-protein interactions, drug-target interactions, and
 CC gene expression profiles. This polynucleotide sequence represents the DNA
 CC encoding a human PMOD protein of the invention
 CC
 SQ Sequence 2572 BP; 780 A; 489 C; 509 G; 794 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 1,09e-146 Length: 2572
 Score: 1385.00 Matches: 267
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 8 Gaps: 0
 US-10-649-273-2_COPY_148_414 (1-267) x ABT23207 (1-2572)

QY	1	MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProheLeuVal	20
Db	585	ATGAGGCGTCATGACACTACTACTAGTTAGCTTGACCAATTAAGTAGATTTCTTTTATGT	644
QY	21	LeuLeuIleSerGlyIglYHisCysLeuLeuAlaLeuValGlnGlyValSerAspHeu	40
Db	645	CTTTTGATTTCTGAGGTCACTGTCTGTGGCATAGTTCAAGAGGTTTCAGATTTTCTG	704
QY	41	LeuLeuGlyIylSerLeuAspIleLeIaPArgIylAspMetLeuAspLysValAlaArg	60
Db	705	CTTCTTGGAAAGTCTTTTGACATACACACAGGTGACATGCTTGGCAAGGTGGCAAGAA	764
QY	61	LeuSerLeuIleIylHisIleProGluCysSerThrMetSerGlyIylValAlaIleGluHis	80
Db	765	CTTTCTTTAATTAACAATCCAGAGTCTCCACATGAGTGGGGAAAGCCATTAAGACT	824
QY	81	LeuAlaIylSGlnGlyAsnArgPheHisIspheAspIleIylsProProLeuHisHisAlaIyls	100
Db	825	TTGGCCAAACAAAGAAATAGATTTTCATTTTGACATCAACCACTCCCTTGATATAGCTAA	884
QY	101	AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysIyls	120
Db	885	AATGTGATTTTCTTTTACTGGACTTCAACACGTTACTGATAAATATATATGAAAG	944
QY	121	GluIylSGlnGluGlyIylIleGluIylGlyGlnIleLeuSerSerAlaAlaAspIleAlaIa	140
Db	945	GAAAAAGGAGAAAGTATGAGAAAGGGCAAAATCTGTCTTCACGACGACACTTCTGCC	1004
QY	141	ThrValGlnHisIsthMetAlaCysHisIleuValIylsArgThrHisArgAlaIleLeuPhe	160
Db	1005	ACAGTAGACGACACAAATGGCATGTCACTTGTGAAABAACACATCGGGCTTATCTGTTT	1064
QY	161	CysIylSGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyIylVal	180
Db	1065	TGTAAGCAGAGAGACTTGTACTTCAAAATATATGACAGTACTGTGTCATCTGTGTGTC	1124
QY	181	AlaSerAsnPheIylIleArgArgAlaLeuGlnIleLeuThrAsnAlaIsthGlnCysThr	200
Db	1125	GCAAGTAATCTTATATATCCGACAGCTCTGAAATTTTAAACAAAGCAACAGATGCACT	1184
QY	201	LeuLeuCysProProProArgLeuCysThrAspAsnGlyIleMetIleAlaIsthAsnGly	220
Db	1185	TTGTTGTGTCTCTCCCTCCAGACTAGCACTGATATATGGCATTTATATTCAGTGAATGT	1244
QY	221	IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGlnGlyIylIleArgIylGlu	240
Db	1245	ATTGAAAGACTACGTCGTGGCTTGGCAATTTATACATGACATGAAAGGATCCGCTATGA	1304
QY	241	ProIylsCysProLeuGlyValAspIleSerLysGluValGlyGlnAlaSerIleLysVal	260
Db	1305	CCAAATATCTCTCTTGAGTAGACATATCAAAAGAGTGGAGAAAGCTTCCATTAAGTA	1364
QY	261	ProGlnLeuLysMetGlnIle 267	
Db	1365	CCACATTAATAAATGAGATA 1385	
RESULT 6			
AAD46856			
ID	AAD46856 standard, cDNA, 1820 BP.		
AC	AAD46856,		
XX	27-JAN-2003 (first entry)		
DT	Human glycoprotease 28472 cDNA.		
DE	Human, adenosine deaminase; seven transmembrane domain receptor; cancer;		
KW	7TM, glycoprotease; immune disorder; Iga deficiency; allergy; arrhythmia;		
KW	rheumatoid arthritis; diabetes; atherosclerosis; cardiovascular disorder;		
KW	hyperension; ischaemic heart disease; obesity; myocardial infarction;		
KW	endothelial cell disorder; Grave's disease; psoriasis; brain disorder;		
KW	Parkinson's disease; Alzheimer's disease; hemotopoietic disorder;		
KW	cerebral oedema; metabolic disorder; liver disorder; platelet disorder;		

KW	chromosome mapping; tissue typing; gene therapy; neuroprotective;
KW	cystostatic; anorectic; cardiant; haemostatic; gene; ss.
XX	
OS	Homo sapiens.
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	146..1390
FT	/tag= "a
FT	/product= "human 28472 protein"
FT	/note= "This region is specifically claimed as SEQ ID NO:
FT	6 in claim 1 of the specification"
PN	
PD	WO200274960-A2.
PD	
PP	26-SEP-2002.
PP	
PR	08-NOV-2001; 2001WO-US051427.
PR	
PR	08-NOV-2000; 2000US-0246768P.
PR	08-NOV-2000; 2000US-0246772P.
PR	15-NOV-2000; 2000US-0249185P.
PA	(MILL-) MILLENNIUM PHARM INC.
PI	
DR	Lelby KR, Kapeller-Libermann R, Glucksmann M;
DR	WPI, 2002-759898/82.
PT	P-PSDB; AAE29234.
PT	New 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid molecules,
PT	useful for diagnosing and treating cancer, immune, cardiovascular,
PT	hematopoietic, brain, pain, metabolic, liver or platelet disorders, and
PT	in pharmacogenomics.
PS	
PS	Claim 1; Fig 8; 178pp; English.
CC	The present invention relates to novel 38650, 28472, 5495, 65507, 81588
CC	or 14354 nucleic acid molecules which encode adenosine deaminase, glyco-
CC	protease or seven transmembrane domain (7TM) receptor family members.
CC	Sequences of the invention are useful in diagnosis and treating cancer
CC	or aberrant cellular proliferation and/or differentiation (e.g. colon or
CC	lung cancer), immune disorders (e.g. selective IGA deficiency, rheumatoid
CC	arthritis, diabetes, allergy), heart or cardiovascular disorders (e.g.,
CC	hypertension, atherosclerosis, arrhythmias, ischemic heart disease,
CC	myocardial infarction, thrombus) including endothelial cell disorders
CC	(e.g. psoriasis, Grave's disease), hematopoietic disorders, brain
CC	disorders (e.g. cerebral oedema, Parkinson's or Alzheimer's disease),
CC	pain and metabolic disorders (e.g. obesity), liver disorders or platelet
CC	disorders. They are also useful in screening assays, predictive medicine
CC	(e.g. diagnostic assays), prognostic assays, monitoring clinical trials
CC	and pharmacogenetics) and prophylactic and therapeutic methods. The
CC	nucleic acids may also be used in chromosome mapping, tissue typing and
CC	forensic biology and as surrogate markers. Sequences of the invention are
CC	also used in gene therapy. The present sequence is human glycoprotease
CC	28472 CDNA
SQ	
SQ	Sequence 1820 BP; 543 A; 365 C; 393 G; 518 T; 0 U; 1 Other;
Alignment Scores:	
Pred. No.:	1,07e-142 Length: 1820
Score:	1348..00 Matches: 260
Percent Similarity:	96.50% Conservative: 3
Best Local Similarity:	97.38% Mismatches: 4
Query Match:	97.33% Indels: 0
DB:	Gaps: 0
US-10-649-273-2_COPY_148_414 (1-267) x AAD46856 (1-1820)	
OY	1 MetGUAAlAHlAAlAleuthrXllleArGLearLLeThAsrLyVeValGlUpheProPhelueVal 20
Db	587 ATGGAGCGCTCAACACTTATTCATTAGCGTTGCCAAATAAGTAATAATTTCTTTTTTAGGT 646
OY	21 LeuauilSeSerGlyClYhiscCysleuLeuAlaleuValGlncIyValSeraphePeLeu 40

[illegible]

```

PT /note= "This CDS is specifically claimed in claim 1"
XX
XX US2003009017-A1.
XX
XX PD
XX PD
XX PD
XX PD
XX PF 08-NOV-2001; 2001US-00012140.
XX
XX PR 08-NOV-2000; 2000US-0246768P.
XX PR 08-NOV-2000; 2000US-0246772P.
XX PR 15-NOV-2000; 2000US-0249185P.
XX
XX PA (LEIBY K R.
XX PA (KAPPELIER-LIBERMANN R.
XX PA (GLUCKSMANN M A.
XX
XX PI Leiby KR, Kapellier-Libermann R, Glucksmann MA;
XX
XX DR WPI: 2003-428888/40.
XX DR P-Patent: ABU09569.
XX
XX CC The invention relates to an isolated 38650 (encoding adenosine
XX CC deaminase), 28472 (encoding a glycoproteinase), 5495 (encoding a 7
XX CC transmembrane receptor), 65507, 81588 or 14354 nucleic acid molecule (or
XX CC a sequence which is at least 60% identical to the six nucleic acids or
XX CC their open reading frames, fragments of at least 15 nucleotides,
XX CC naturally occurring variants, or a DNA insert of the plasmid deposited
XX CC with the American Type Culture Collection as Accession No. not defined in
XX CC the specification, which encodes the amino acid sequence). Also included
XX CC are a host cell containing the nucleic acids (used to produce the
XX CC protein), the encoded protein, an antibody that selectively binds to
XX CC the polypeptide, and identifying a compound that binds to/modulates the
XX CC activity of the polypeptide. The nucleic acid molecules, polypeptides and
XX CC methods are useful for diagnosing, treating cancer, aberrant cellular
XX CC cardiovascular and/or differentiation, immune disorders, heart disorders,
XX CC haematopoietic disorders, blood vessel disorders, brain disorders, pain
XX CC and metabolic disorders, liver disorders and platelet disorders (many
XX CC examples of these disorders are given in the specification). The present
XX CC sequence is the Human cDNA 28472 encoding a glycoproteinase
XX
SQ Sequence 1820 BP; 543 A; 365 C; 393 G; 518 T; 0 U; 1 Other;
SO
Alignment Scores:
Pred. No.: 1.07e-142 Length: 1820
Score: 1348.00 Matches: 260
Percent Similarity: 98.50% Conservative: 3
Best Local Similarity: 97.38% Mismatches: 4
Query Match: 97.33% Indels: 0
DB: 8 Gaps: 0
US-10-649-273-2_COPY_148_414 (1-267) x ACA60887 (1-1820)
OY 1 MetGluAlaHisLeuThrIleArgLeuThrAsnValGlnProPheLeuVal 20
Db 587 ATGGAGGCTCATGCACCTTAATTTAGTGTGACCATAAAGTAAATTTCCCTTTTAGATT 646
OY 21 LeuLeuIleSerGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
Db 647 CTTCGATTTCTGGAGGTCACCTGCTCTTGCGACTTAGTTCAAGAGTTTCAAGATTTCG 706
OY 41 LeuLeuGlyIleValSerLeuAsnTrpIleAlaProGlyValSerMetLeuAspLeuVal 50
Db 707 CTTCGTAAGAAGCTTTTGACATAGCACCAAGGGAATGCTTGACAACAAGTGGCAAGAAGA 766
OY 61 LeuSerLeuIleLeuHisIleProGlyCysSerTrpMetSerCylValValIleGlnHis 80

```

```
DB 767 CTTCTTTAATAAATCCAGATGCTCCACATGATGCTGGAAAGCATAGAACAT 826
QY 1LeuAlaIySGInGlyAsnArgPheHisPheAspIleYsProPheUhiSHISAlaIys 100
DB 827 TTGGCCAAACAAGGAATAGATTTCATTTTGATCAATCAACTCCCTTGATCATGCTAA 886
QY 101 AenCysAspPheSerPheThrGlyLeuGlnHisValThraSplyIleIleMetLysLys 120
DB 887 AATTGGATTTTCTTTTACTGATCGACTCAACAGTTACTGATTAATAATGAAACAG 946
QY 121 GlnYsGlnGlyGlyIleGlnYsGlyGlnIleLeuSerSerIleAlaAspIleAla 140
DB 947 AACACAGAGAGAAAGTATGAGAAAGGAGCAATCTGCTTCACACAGCATTCGCTGCC 1006
QY 141 ThrValGlnHisThrMetAlaCysHisLeuValIysArgThrHisArgAlaIleLeuPhe 160
DB 1007 AAGGTACAGACACATGCGATGTCATCTTTGTAAGAAAGAACATCGGCTATTCGTTT 1066
QY 161 CysLysGlnArgAspLeuProGlnAsnAsnAlaValIleuValAlaSerGlyVal 180
DB 1067 TGTAAAGCAGAGACTGTTTACTCAAAATATGACGATCTGGTTCATCGGCGGTGC 1126
QY 181 AlaSerAspPheTyrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr 200
DB 1127 GGAAGTAACTTCATATTCGACAGAGCTCTGGAATTTTAAACAAACGACACAGTGCAC 1186
QY 201 LeuLeuCysProProArgLeuLeuCysThrAspAspGlyIleMetIleAlaTPasGly 220
DB 1187 TTGTTGTGTCCTCTCCACAGCATGACATGATTAATGGCAATTAATGATGATGGAATGGT 1246
QY 221 IleguArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGlyIleArgYrgIu 240
DB 1247 ATTGAAAGCTAGCTGCTGGCTTGGCATTTTACATGACATAGAGCAATCCCTATGAA 1306
QY 241 ProLysCysProLeuGlyValAspIleSerLysGlyValGlyAlaSerIleLysVal 260
DB 1307 CCAAAATGCTCTTTCGAGTACGACATATCAAAAGAGTTGAGAGAGCTTCATTAAGA 1366
QY 261 ProGlnLeuLeuMetGlnIle 267
DB 1367 CCACATTAATAAATGAGATA 1387
RESULT 8
ABS57020
ID ABS57020 standard; cDNA; 1821 BP.
AC ABS57020;
XX 30-JAN-2003 (first entry)
DE cDNA encoding novel human glycoprotease 28472.
XX
XX Cancers; aberrant cell proliferation; aberrant cell differentiation;
XX breast cancer; ovarian cancer; prostate cancer; colon cancer;
XX lung cancer; immune disorder; heart disorder; cardiovascular disorder;
XX endothelial disorder; hematopoietic disorder; blood vessel disorder;
XX brain disorder; pain; metabolic disorder; liver disorder; diabetes;
XX platelet disorder; carcinoma; sarcoma; leukemia; Hodgkin's disease;
XX autoimmune disorder; hypertension; atherosclerosis; heart failure;
XX myocardial infarction; ischemic heart disease; Crohn's disease;
XX Grave's disease; Kawasaki syndrome; Raynaud's disease; aneurysm;
XX cerebral ischemia; peripheral neuropathy; Alzheimer's disease;
XX Parkinson's disease; anorexia nervosa; cachexia; glycoprotease; gene; ss.
OS Homo sapiens.
XX
XX Key location/Qualifiers
XX CDS 147..1391
XX /tag= a
XX /product= "glycoprotease 28472"
XX /note= "Specifically claimed in claim 1"
```

```
PN K0200277233-A2.
XX
XX 03-OCT-2002.
XX
XX 08-NOV-2001; 2001WO-US046724.
XX
XX 08-NOV-2000; 2000US-0246768P.
XX
XX 08-NOV-2000; 2000US-0246772P.
XX
XX 15-NOV-2000; 2000US-0249185P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
XX Leiby KR, Kapeller-Libermann R, Glucksmann M;
XX
XX WPI; 2003-029938/02.
XX
XX P-PSDB; ABG71162.
XX
XX New adenosine deaminase, glycoprotease and seven transmembrane domain
XX nucleic acids and polypeptides, designated 38650, 28472, 5495, 65507,
XX 81588 and 14354, useful for treating e.g. leukemias, Hodgkin's disease or
XX hypertension.
XX
XX Claim 2; Fig 8A-B; 178pp; English.
XX
XX The invention describes isolated 38650, 28472, 5495, 65507, 81588 and
XX 14354 nucleic acid molecules (I) and their encoded polypeptides (II). The
XX 38650 nucleic acid molecule comprises a sequence encoding adenosine
XX deaminase. The 28472 nucleic acid molecule comprises a sequence encoding
XX a human glycoprotease. The 5495, 65507, 81588 and 14354 comprise
XX sequences that encode a human seven transmembrane domain (7TM). The
XX 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid and polypeptide
XX sequences are useful for diagnosing, preventing or treating a subject
XX with or at risk of developing a disorder, e.g. cancer or aberrant
XX cellular proliferation and/or differentiation (e.g. breast, ovarian,
XX prostate, colon or lung cancer), immune disorders, heart disorders,
XX cardiovascular disorders, endothelial disorders, hematopoietic disorders,
XX blood vessel disorders, brain disorders, pain and metabolic disorders,
XX liver disorders or platelet disorders. These disorders include carcinoma,
XX sarcoma, leukemia, Hodgkin's disease, autoimmune disorders,
XX hypertension, atherosclerosis, heart failure, myocardial infarction,
XX ischaemic heart disease, Crohn's disease, Grave's disease, Kawasaki
XX syndrome, Raynaud's disease, aneurysm, cerebral ischemia, peripheral
XX neuropathy, Alzheimer's disease, Parkinson's disease, anorexia nervosa,
XX cachexia or diabetes. This sequence encodes the novel human glycoprotease
XX 28472
XX
XX SQ Sequence 1821 BP; 543 A; 365 C; 394 G; 518 T; 0 U; 1 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.07e-142 Length: 1821
XX Score: 1348.00 Matches: 260
XX Percent Similarity: 98.50% Conservative: 3
XX Best Local Similarity: 97.38% Mismatches: 4
XX Query Match: 97.33% Indels: 0
XX DB: Gaps: 0
XX
XX US-10-649-273-2_COPY_148_414 (1-267) x ABS57020 (1-1821)
QY 1 MetGlnAlaHisAlaLeuThrIleArgLeuThrAsnLysValGlnPheProPheLeuVal 20
DB 588 ATGAGGCTCATGCACTTACTATTAAGTACCAATTAAGTAAGATTTCTTTTATGTT 647
QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
DB 648 CTTTGAATTTCTGAGAGTCACTGCTTGGCATTAAGTTCAGAGATTTCTG 707
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
DB 708 CTTCTTGAAGAGCTTTGGACATRGACACAGGTACATGCTTGACAAAGTGGCAAGAGA 767
QY 61 LeuSerLeuIleLysHisProGlnCysSerThrMetSerGlyGlyLysAlaIleGlnHis 80
DB 768 CTTTCTTTAATAAATCCAGATGCTCCACATGATGATGAGTGGAAAGCATAGAACAT 827
```

QY 81 LeuAlaIySGInGlyAsnaIrgPhehiSphEaspIleIySfProleuhiSihIaIaIyS 100
 Db 828 TTGGCCAAACAAAGAAATGATTTGATTTGACATCAAACTCCCTTGATCATGTGAAA 887
 QY 101 AenCySaSpPheSerPheThrgIyLeuGInhiSValIThrAspIySileIleMeIySlyS 120
 Db 888 AATTGTGATTTTCTTTTACTGTGACCTTCAACAGTTACTGATTAATAATATGAAGAACG 947
 QY 121 GIuIySGInGluGlyIleGIuIySGInGInIleuSerSerIaIaIaSpIleIaIaIa 140
 Db 948 AAACAGAGAGAAAGATTAAGAAAGGCGCAATCTGTCTTCAGACAGACATTTGCTGCC 1007
 QY 141 ThrValGInhiSthrmEAlaCySihIaIeValIySArGthrhIaIaIaIleuPhe 160
 Db 1008 ACAGTACAGACACATGCGATCTCATCTTGTGAAAAGAACACATCGGCTATTCTGTTT 1067
 QY 161 CySlySGInArGaSpIleuProGInaSpnaIaValIeValaIaIaSerGIyIyVal 180
 Db 1068 TGTAAAGCAGAGACTTGTACTCAATAATATGACAGTACTGTGATCTGTGCTGCTG 1127
 QY 181 AlaSerAspPheThrgIleArGaIaIeGInIleuThrhAsnaIaIaThrhInCySthr 200
 Db 1128 GCAAGTAACTTCTATATCCGACAGCTCTGGAATAATTTAAACAACGACACAGTGCAT 1187
 QY 201 LeuLeuCySfProProArGleuCySthrhAspnaGIyIleMeIleIaIaTPaSnGIy 220
 Db 1188 TTGTGTGTCTCTCTCCAGACTATGACATGATTAAGCATTTATGATTCATGGAATGCT 1247
 QY 221 IleGIuArGIeUArGaIaGIyLeuGIyIleUhiSAspIleGIuGIyIleArGIyGlu 240
 Db 1248 ATTGAAGAAGACTAGCTCTGCTGGCATTTTATCATGACATGAAGGCATCCGCTATGAA 1307
 QY 241 ProIyCySfProleuGIyValaIaSpIleSerIySGInValaIyGIuIaIaSerIleIyVal 260
 Db 1308 CCAAAATGTCTCTTGAGATGACATATCAAAAGAGTTGAGAGACTTCATTAAGATA 1367
 QY 261 ProGInleuIyMeIeGIuIle 267
 Db 1368 CCACATTAATAATGAGATTA 1388
 Db
 RESULT 9
 ADA52832
 ID ADA52832 standard; cDNA; 2208 bp.
 AC ADA52832;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human coding sequence, SEQ ID 400.
 XX
 KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Neotropic;
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
 KW inflammatory disease; osteoporosis; neurological disease; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1293569-A2.
 XX
 PD 19-MAR-2003.
 XX
 PF 21-MAR-2002; 2002EP-00006586.
 XX
 PR 14-SEP-2001; 2001JP-00328381.
 XX
 PR 24-JAN-2002; 2002US-0350435P.
 XX
 PA (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Iehi S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
 XX

DR WPI; 2003-395539/38.
 DR P-PSDB; ADA54471.
 XX
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 XX
 PS Claim 1; SEQ ID NO 400; 205pp; English.
 XX
 CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.
 XX
 SO Sequence 2208 bp; 660 A; 454 C; 451 G; 643 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3.3e-126 Length: 2208
 Score: 1204.00 Matches: 239
 Percent Similarity: 89.51% Conservative: 0
 Best Local Similarity: 89.51% Mismatches: 4
 Query Match: 86.93% Indels: 24
 DB: 10 Gaps: 1
 US-10-649-273-2_COPY_148_414 (1-267) x ADA52832 (1-2208)
 QY 1 MetGIuIaIhIaIaIeUThrhIeArGleuThrhAsnIySValaIaIaPheProPheUVal 20
 Db 785 ATGAGAGCTCAATGACCTTACTATTAAGTTGACCAATTAAGTAATTTCTTTTATGAT 844
 QY 21 LeuLeuIleSerGIyGIyhiSfSeIeUeUaIaIeUValaIaIaGIyIyValaIaIaIa 40
 Db 845 CTTTGTGATTTTGTGAGAGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 904
 QY 41 LeuLeuGIySfSerIeUaSpIleAlaProGIyAspMeIleUaSpIyValaIaIaIaIa 60
 Db 905 CTTCTGGAAGAGCTCTTGGACATGACACAGGTGACATGCTTGAACAAGGTGGCAAGAGA 964
 QY 61 LeuSerIeUleIySfProGIyCySfSerThrhMeIeSerGIyGIyValaIaIaIaIa 80
 Db 965 CTTCTTAAATTAATAACATCCAGAGTCTCCACCAATGAGTGGGGAAGCCATGACAT 1024
 QY 81 LeuAlaIySGInGlyAsnaIrgPhehiSphEaspIleIySfProleuhiSihIaIaIyS 100
 Db 1025 TTGGCCAAACAAAGAAATGATTTGATTTGACATCAAACTCCCTTGATCATGTGAAA 1084
 QY 101 AenCySaSpPheSerPheThrgIyLeuGInhiSValIThrAspIySileIleMeIySlyS 120
 Db 1085 AATTGTGATTTTCTTTTACTGTGACCTTCAACAGCTTACTGATTAATAATATGAAGAACG 1144
 QY 121 GIuIySGInGluGlyIleGIuIySGInGInIleuSerSerIaIaIaSpIleIaIaIa 140
 Db 1145 GAATAAGAAGAAAGATTAAGAAAGGCGCAATCTGTCTTCAGACAGACATTTGCTGCC 1204
 QY 141 ThrValGInhiSthrmEAlaCySihIaIeValIySArGthrhIaIaIaIleuPhe 160
 Db 1205 ACAGTACAGACACATGCGATCTCATCTTGTGAAAAGAACACATCGGCTATTCTGTTT 1264
 QY 161 CySlySGInArGaSpIleuProGInaSpnaIaValIeValaIaIaSerGIyIyVal 180
 Db 1265 TGTAAAGCAGAGACTTGTACTCAATAATATGACAGTACTGTGATCTGTGCTGCTG 1324
 QY 181 AlaSerAspPheThrgIleArGaIaIeGInIleuThrhAsnaIaIaThrhInCySthr 200
 Db 1325 GCAAGTAACTTCTGATTCGCGAGCTCTGGAATTTTAAACAAGCACAACAGTGCAT 1384
 QY 201 LeuLeuCySfProProArGleuCySthrhAspnaGIyIleMeIleIaIaTPaSnGIy 220
 Db 1385 TTGTGTGTCTCTCTCCAGACTATGACATGATTAATGATATGATGCA----- 1435
 QY 221 IleGIuArGIeUArGaIaGIyLeuGIyIleUhiSAspIleGIuGIyIleArGIyGlu 240

DB 1435 ----- 1435
QY 241 ProlyscysProleuGlyValAspIleSerlyGluValGlyGluAlaSerIlelyVal 260
DB 1436 ---TGATGTCCTTGGAGTACATATCAAAAGAGTTGGAGAGCTTCATTAAGTA 1492
QY 261 ProGlnLeuLysMetGluIle 267
DB 1493 CCACAAATTAATAATGAGATA 1513
RESULT 10
ADQ24627
ID ADQ24627 standard; DNA; 2890 BP.
XX
AC ADQ24627;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 7447.
XX
KM soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 7447; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual,
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 2890 BP; 869 A; 609 C; 611 G; 789 T; 0 U; 12 Other;
XX
Alignment Scores:
Pred. No.: 4,83e-126 Length: 2890
Score: 1204.00 Matches: 239
Percent Similarity: 89.51% Conserves: 0
Best Local Similarity: 89.51% Mismatches: 4
Query Match: 86.93% Indels: 24
DB: 12 Gaps: 1
US-10-649-273-2_COPY_148_414 (1-267) x ADQ24627 (1-2890)
QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20

DB 1442 ATGAGGCTCATGACTTACTATTAGGTGACCAATTAAGTAGAATTTCTTTTACTT 1501
QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuValLeuValGlnGlyValAspPheLeu 40
DB 1502 CTTTGAATTTCTGGAGGCTCACTGCTGTGGCACTTAGTTCAGAGATTTCGATTTCG 1561
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
DB 1562 CTTCCTGAAAGTCTTTGGACATAGCACACAGGTGACATGCTTGACCAAGGTGGCAAGAGA 1621
QY 61 LeuSerLeuIleLysHisProGluCysSerThrMetSerGlyLysValAlaIleGlnHis 80
DB 1622 CTTCCTTAATTAACATCCAGAGTGTCCACCATGAGTGGGAAAGCCATAGAAAT 1681
QY 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLys 100
DB 1682 TTGGCCAAACAGAAATAGATTCAATTTGACATCAAACTCCCTTCACATAGCTTAA 1741
QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
DB 1742 AATTGATTTTCTTTTACTGACCTTCAACAGTTACTGATTAATATATAGAAAAAG 1801
QY 121 GluLysGluGluGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 140
DB 1802 GAAAAAGAGAGAGTATTGAGAAAGGCAATCTGCTTCAGACGACGACATTGCTGCC 1861
QY 141 ThrValGlnHisSerThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
DB 1862 ACAGTACGACACAAAGGCAATGTCATCTTGTAAGAAAGAACACATCGGGCTAATCTGTT 1921
QY 161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyVal 180
DB 1922 TGTAGCGAGAGAGCTTGTATCTCAAAATATAGCAAGTGTGATCTGTGTGTCTC 1981
QY 181 AlaSerAsnPheThrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThr 200
DB 1982 GCAGTAACTTCTGTATCCGAGAGCTGTGAAATTTTAAACAAAGCAACAGTGCCT 2041
QY 201 LeuLeuCysProProProAlaGluCysThrAspAsnGlyIleMetIleAlaIlePheAsnGly 220
DB 2042 TTGTTGTGTCCTCTCCAGCATATGCACTATATAGCATTAATGATTGCA----- 2092
QY 221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGlu 240
DB 2092 ----- 2092
QY 241 ProlyscysProleuGlyValAspIleSerlyGluValGlyGluAlaSerIlelyVal 260
DB 2093 ---TGATGTCCTTGGAGTACATATCAAAAGAGTTGGAGAGCTTCATTAAGTA 2149
QY 261 ProGlnLeuLysMetGluIle 267
DB 2150 CCACAAATTAATAATGAGATA 2170
RESULT 11
ADE31345/C
ID ADE31345 standard; DNA; 3358 BP.
XX
AC ADE31345;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human diagnostic and therapeutic polynucleotide (dithp), SEQ ID No 100.
XX
DB
XX
KM diagnostic and therapeutic polynucleotide; dithp; antiarteriosclerotic;
KM antiinflammatory; cerebroprotective; antilipidemic; antidiabetic;
KM immunosuppressive; neuroprotective; nootropic; neuroleptic; tranquilizer;
KM osteopathic; antiarthritic; antirheumatic; cytostatic; hepatotropic;
KM virocidic; haemostatic; anti-HIV; antithyroid; chytromimetic;
KM dermatological; antibacterial; fungicide; antiparasitic; anticonvulsant;
KM thrombolytic; anticoagulant; anorectic; vasotropic; antidiacer;
KM gene therapy; protein replacement therapy; human; gene; ds.

Db 804 GAAAAAGAGAGGACATTGAGAGGGGCAATCTGTCTACGCTGACGACATTCCTGCT 863
Qy 141 ThrValGlnHisThrMetAlaCysHisLeuValIleArgThrHisArgAlaIleLeuPhe 160
Db 864 GGGGTACACATCCACACAGGCTCCACCTTCCGAAAAGAACACATCGCGCTATTCCTTT 923
Qy 161 CysIleGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyVal 180
Db 924 TCAGACGCAAAAATTTGGCTCTCTCCAGCTAACGCAAGTATAGTTGATCTGGAGGCTT 983
Qy 181 AlaSerAspPheValIleArgAlaIleGlnIleLeuThrAsnAlaThrGlnCysThr 200
Db 984 GGAGTAACTTGTACATCCGAAAAGCATTTGCAAAATTCGCAATGCAACGCATGACCG 1043
Qy 201 LeuLeuCysProProArgLeuCysThrAspAsnGlyIleMetIleAlaTPAsnGly 220
Db 1044 TTGTTGTGTCCTCT-TCAGACGCTGCTGACATGCGCATGATGATTCATGGAATGGA 1102
Qy 221 IleGluArgLeuArgAlaGlyLeuGlyIleLeu-HisAsp-IleGluGlyIleArgTyrG 240
Db 1103 ATTGGAAGATTACGTGCTGCTGCTTACCATGATGATGAGACATTCGGTATT 1162
Qy 240 IupProIleCysProLeuGlyVal--AspIleSerIleGlyValGlyGluAlaSerIle 259
Db 1163 AACCCAAATGTCCTCTTGAAGTGAAGCATTCGCCGAAGAGTTGGCAGA---AGCTTGC 1219
Qy 259 yValProGlnLeu 263
Db 1220 CCATTAAAAAGTTA 1233
RESULT 13
ID AAS84622 standard; cDNA; 2734 BP.
XX AAS84622;
AC AAS84622;
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #20426.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001MO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
XX P-PSDB; ABG20435.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID NO 20426; 103bp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2734 BP; 773 A; 545 C; 639 G; 763 T; 0 U; 14 Other;
Alignment Scores:
Pred. No.: 2,61e-57 Length: 2734
Score: 599.00 Matches: 131
Percent Similarity: 71.57% Conservative: 10
Best Local Similarity: 66.50% Mismatches: 22
Query Match: 43.25% Indels: 34
Gaps: 6
US-10-649-273-2_COPY_148_414 (1-267) x AAS84622 (1-2734)
Qy 71 ThrMetSerGlyIleValAlaIleGlnIleLeuAlaIleGlnIleValArgPheHisPhe 90
Db 206 ACCCTGCTAGGGGGGCGCGCT-----AGATTTCAT--- 235.
Qy 91 AspIleIleProProLeuHisIleAlaIleAsnGly----- 102
Db 236 -----CCTATTCTCCAGATGAAGATGACGCTGTAGAACCTGCGACGA 283
Qy 103 -----AspPheSerPheThrGlyLeuGlnHisValThrAsp 114
Db 284 TCTGCTTCTGGGAGAGCCGACGAGCTTTTACTATGCGGAGCAATTCGGAGTACG 343
Qy 115 LysIleIleMetIleValGlyIleGluIleGluIleGluIleGluIleLeu 132
Db 344 CGCTTACATGCGACGACGACGACGACGAGGGGGGATGAGAGAGGGGCAATCTG 403
Qy 133 SerSerAlaAlaAspIleAlaIleAlaIleThrValGlnHisIleThrMetAlaCysHisLeuValIle 152
Db 404 TCTTCAGCAGCAGCATTGCTGCCACAGTACGACACACAAATGCGATCTTGTGAAA 463
Qy 153 ArgThrHisArgAlaIleLeuPheCysIleArgIleAspLeuProGlnAsnAsnAla 172
Db 464 AGAACACATCGGGCTATTCTGTTTGTGAAGAGAGAGCTTTTACTCAAAATTAATGCA 523
Qy 173 ValIleuValAlaSerGlyIleValAlaSerAspPheValIleArgArgAlaLeuGluIle 192
Db 524 GTACTGTTGATCTGCGTGTGCTGCCAGTAACTTCTATATCCGACAGCTTGGAATT 583
Qy 193 LeuThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLeuCysThrAspAsn 212
Db 584 TTACAAAGCAGAACAGTGCACCTTGTGTCTCTCCGACGACGATGACGATGATAT 643
Qy 213 GlyIleMetIleAlaTPAsnGlyIleGluArgLeuArgAlaGlyIleGluIleIleuHis 232
Db 644 GGCATTATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 703
Qy 233 AspIleGluGlyIleArgTyrGluProIle-----CysProLeuGlyVal 247
Db 704 GACATGAGAGGCGATCGCTATGACCAAGATATGTCGCTCTTTCAGAGGCTG 754
RESULT 14
ID ADL86725 standard; DNA; 371 BP.
XX

```
AC ADL86725;
DF 20-MAY-2004 (first entry)
DE DNA up-regulated in murine multipotent progenitor cells Segid 3118.
XX gene potential; multi-lineage; cell commitment; haematopoietic stem cell;
KM HSC; multipotent progenitor; MPP; common lymphoid progenitor; CLP;
XX common myeloid progenitor; CMP; bone marrow stem cell; mouse; murine; ds.
OS Mus sp.
XX WO2003093445-A2.
PN 13-NOV-2003.
PD 13-NOV-2003.
XX PF 05-MAY-2003; 2003WO-US014114.
PR 03-MAY-2002; 2002US-0377383P.
PA (STOW-) STOWERS INST MEDICAL RES.
XX Li L;
XX MPI; 2004-022656/02.
DR Classifying an unknown multi-lineage affiliated gene comprises isolating
PT expressed nucleic acid sequences from the discrete cell sub-populations.
XX Claim 8; SEQ ID NO 3118; 123bp; English.
XX This invention relates to a novel method for predicting gene potential by
CC associating nucleic acid sequences of unknown function with particular
CC sub-population profiles. Specifically, it refers to classifying an
CC unknown multi-lineage affiliated gene by collecting hybridisation data to
CC develop a gene expression map, in order to determine the discrete sub-
CC population where it is expressed. The present invention describes methods
CC for predicting the lineage commitment of genes associated with the self-
CC renewing haematopoietic (blood) stem cells (HSCs), as well as the non-
CC self renewing multipotent progenitors (MPPs), common lymphoid progenitors
CC (CLPs) and common myeloid progenitors (CMPs), which are collectively
CC referred to as bone marrow stem cells populations. As such, these methods
CC can be used to identify associated multi-lineage affiliated genes and
CC hence the underlying molecular mechanisms in physiological haematopoietic
CC development. This polynucleotide sequence is DNA associated with a murine
CC MPP sub population of cells of the invention.
SQ Sequence 371 BP; 104 A; 75 C; 84 G; 92 T; 0 U; 16 Other;
XX
Alignment Scores:
Pred. No.: 1.18e-43 Length: 371
Score: 468.00 Matches: 92
Percent Similarity: 78.86% Conservative: 5
Best Local Similarity: 74.80% Mismatches: 26
Query Match: 33.79% Indels: 0
DB: 12 Gaps: 0
US-10-649-273-2_COPY_148_414 (1-267) x ADL86725 (1-371)
OY AspIlealaaIathrValgInhiStrMetAlaCyHiSleueVallyAargThrhISArg 156
Db | ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2 GAGCATGTCGTCTGGCGGTACAGCAATGCACAACGCCGTGCACCTTCCGAAMAAACACACATC 61
OY 157 AlaleluenheCyselySGlnarGaSpLeuleUProGlinaSnasnaIlavalleuValala 176
Db | ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
62 GCTATTCTGTTTGACAGAAGCAAGAAATTGCTCTCCACGCTAACACCAGTAATTAGTTGTA 121
OY 177 SerGIyglYvalaIsSerasnPheryllEargaxAlaleuGluileutrnAsnla 196
Db | ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
122 TCtGcAgGgtGtgTtcCaAGTAactTGtAcAtccCAaMaagCatTTgGaMAatYttgcGAaATGca 181
OY 197 ThrGIncySthrleuLeucycFProBorPrCoArGuLeucyethrsapngnyllemellle 216
```

```

Db      182   ACGCAGTGTACAGTTGTTTGTCGCCACTCCAGACATGTGCACGACGACAAATGGCATCAAGATT 241
Oy      217   ALATTPAENGYYLIEGLUARGHUAARGHLAAGLYLEUGLYYILLEUHISASPILEGJUGLY 236
Db      242   GCATGTGAATGGAATGTGAANGAATTAAGTCGTGCTGNGCTTNGCNTNTTACNTGANTNATAAGAAGAC 301
Oy      237   ILeatGYrYGluPProLyScyeProLeuglYValAsPIleSerLySGluVaIglYGuLaA 256
                |||::|||
Db      302   ATCCNNHTATGNACCBAABATNTCNTMTTGTGAGTAGACATMTTCCAGAGAAGTTGCAGAGCT 361
Oy      257   SerileLyS 259
                ::|||
Db      362   GCCATPAAAA 370

RESULT 15
ADL86726
ID      ADL86726 standard; DNA; 371 BP.
XX      ADL86726;
AC      ADL86726;
XX      ADL86726;
DT      20-MAY-2004 (first entry)
DE      DNA up-regulated in murine multipotent progenitor cells Segid 3119.
DX      gene potential; multi-lineage; cell commitment; haematopoietic stem cell;
KW      HSC; multipotent progenitor; MPP; common lymphoid progenitor; CLP;
KV      common myeloid progenitor; CMP; bone marrow stem cell; mouse; murine; ds.
XX      Mus sp.
OS      WO2003093445-A2.
PN      13-NOV-2003.
PD      05-MAY-2003; 2003WO-US014114.
PF      03-MAY-2002; 2002US-0377383P.
PR      (STOW-) STOMERS INST MEDICAL RES.
XX      PA
XX      Li L;
DR      WPI; 2004-022656/02.
DX      Classifying an unknown multi-lineage affiliated gene comprises isolating
PT      expressed nucleic acid sequences from the discrete cell sub-populations.
XX      Claim 8; SEQ ID NO 3119; 123bp; English.
XX      This invention relates to a novel method for predicting gene potential by
CC      associating nucleic acid sequences of unknown function with particular
CC      sub-population profiles. Specifically, it refers to classifying an
CC      unknown multi-lineage affiliated gene by collecting hybridisation data to
CC      develop a gene expression map, in order to determine the discrete sub-
CC      population where it is expressed. The present invention describes methods
CC      for predicting the lineage commitment of genes associated with the self-
CC      renewing haematopoietic (blood) stem cells (HSCs), as well as the non-
CC      self renewing multipotent progenitors (MPPs), common lymphoid progenitors
CC      (CLPs) and common myeloid progenitors (CMPs), which are collectively
CC      referred to as bone marrow stem cells populations. As such, these methods
CC      can be used to identify associated multi-lineage affiliated genes and
CC      hence the underlying molecular mechanisms in physiological haematopoietic
CC      development. This polynucleotide sequence is DNA associated with a murine
CC      MPP sub population of cells of the invention.
XX      SQ
SQ      Sequence 371 BP; 104 A; 75 C; 84 G; 92 T; 0 U; 16 Other;

Alignment Scores:
Pred. No.:          1,186--43           length:       371
Score:              468.00             Matches:        92
Percent Similarity: 78.86%             Conservative:    5
Best Local Similarity: 74.80%           Mismatches:     26
Query Match:        33.79%             Indels:         0

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 16, 2005, 18:30:49 ; Search time 132.629 Seconds
(without alignments)
3294.036 Million cell updates/sec

Title: US-10-649-273-2_COPY_148_414

Perfect score: 1385
Sequence: 1 MEAHALTIRLNKVFEPFLV.....DISKEVGASIKVPOLKMEI 267

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10649273/runat_15062005_111417_6057/app_query.fasta_1.1429
-DB=Issued Patents NA -QFMT=faaetap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10649273@cgn2_1.1.177 @runat_15062005_111417_6057 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfile1.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1385	100.0	1416	US-09-774-528-177	Sequence 177, App
2	1385	100.0	1526	US-10-067-443-23	Sequence 23, Appl
3	1385	100.0	2197	US-10-067-443-1	Sequence 21, Appl
4	1358.5	98.1	1387	US-10-067-443-21	Sequence 20, Appl
5	824.5	59.5	14364	US-10-067-443-20	Sequence 38, Appl
6	308.5	22.3	94750	US-09-596-002-38	Sequence 806, App
7	307	22.2	1053	US-09-540-236-806	Sequence 884, App
8	291.5	21.0	1059	US-09-252-991A-884	Sequence 801, Appl
9	291.5	21.0	1206	US-09-252-991A-801	Sequence 1, Appl
10	289.5	20.9	1830121	US-09-557-884-1	Sequence 1, Appl
11	289.5	20.9	1830121	US-09-643-990A-1	Sequence 1, Appl
12	282	20.4	996	US-09-902-540-6612	Sequence 6612, Ap

C 13	282	20.4	2582	4	US-09-902-540-503	Sequence 503, App
14	280.5	20.3	1074	4	US-09-543-681A-2341	Sequence 2341, Ap
15	270.5	19.5	1315	1	US-08-087-797-1	Sequence 1, Appl
16	268.5	19.4	1032	4	US-09-489-039A-2050	Sequence 2050, Ap
17	261	18.8	1092	4	US-09-107-532A-2955	Sequence 2955, Ap
18	254	18.3	1008	3	US-08-987-121A-5	Sequence 5, Appl
19	254	18.3	1011	3	US-08-987-121A-3	Sequence 3, Appl
20	250	18.1	1006	3	US-08-961-083-51	Sequence 51, Appl
21	250	18.1	1006	4	US-09-536-784-51	Sequence 51, Appl
22	250	18.1	10974	3	US-08-961-527-214	Sequence 214, App
23	248	17.9	1011	3	US-09-066-512-1	Sequence 1, Appl
24	246	17.8	1011	4	US-09-583-110-2196	Sequence 2196, Ap
25	241	17.4	1011	4	US-09-107-433-1618	Sequence 1618, Ap
26	236.5	17.1	640681	4	US-09-790-988-1	Sequence 1, Appl
27	236	17.0	8822	4	US-09-710-279-727	Sequence 727, App
28	236	17.0	1101	4	US-09-134-000C-1551	Sequence 1551, Ap
29	236	17.0	3993	4	US-09-710-279-3985	Sequence 3985, Ap
30	233	16.8	1107	3	US-09-134-001C-1072	Sequence 1072, Ap
31	220	15.9	1026	3	US-09-149-624-1	Sequence 1, Appl
32	218	15.7	3064	3	US-09-221-017B-794	Sequence 794, App
33	218	15.7	1230230	4	US-09-198-452A-1	Sequence 1, Appl
34	218	15.7	1230230	4	US-09-438-185A-1	Sequence 1, Appl
35	216	15.6	15249	4	US-08-956-171E-102	Sequence 102, App
36	216	15.6	15249	4	US-08-781-966A-102	Sequence 102, App
37	215.5	15.6	4403765	3	US-09-103-840A-2	Sequence 2, Appl
38	215.5	15.6	4411529	3	US-09-602-777A-147	Sequence 147, App
39	211.5	15.3	1155	4	US-09-602-777A-147	Sequence 1, Appl
40	207.5	15.0	1664976	4	US-08-916-421B-1	Sequence 1, Appl
41	207.5	15.0	1664976	4	US-09-692-570-1	Sequence 6, Appl
42	192.5	13.9	1663	4	US-09-328-352-261	Sequence 261, App
43	185.5	13.4	432	4	US-08-311-731A-131	Sequence 131, App
44	181.5	13.1	432	4	US-09-328-352-261	Sequence 261, App
45	169	12.2	876	4	US-09-724-623-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1

US-09-774-528-177
; Sequence 177, Application US/09774528
; Patent No. 6743619

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Goodrich, Ryle
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyun
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Yang, Yonghong
APPLICANT: Xue, Aidong J.
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Wang, Dunrui
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6743619el Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT FILING DATE: 2001-01-30
CURRENT APPLICATION NUMBER: US/09/774,528
NUMBER OF SEQ ID NOS: 441
SOFTWARE: pt_fl_genes Version 2.0
SEQ ID NO 177
LENGTH: 1416
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (205)..(1305)
US-09-774-528-177

Alignment Scores:


```
QY 241 ProlyscysProleuglyValaspilserlysglyValglyValasertilelyVal 260
DB 721 CCAAAATGCTCTTGAGAGTACATATCAAAAGAGTGGAGAGCTTCCATAAAGTA 780
QY 261 ProglinleuylMetglulle 267
DB 781 CCACAATTAAAAATGAGATA 801

RESULT 3
US-10-067-443-1
; Sequence 1, Application US/10067443
; Patent No. 6642041
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE HIGHLY EXPRESSED IN
; TITLE OF INVENTION: SPINAL CORD, MP-1
; FILE REFERENCE: D0073 NP
; CURRENT APPLICATION NUMBER: US/10/067,443
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 2197
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (231)..(1472)
US-10-067-443-1

Alignment Scores:
Pred. No.: 5,23e-177 Length: 2197
Score: 1385.00 Matches: 267
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x US-10-067-443-1 (1-2197)
QY 1 MetglualahisAlaLeuThrIleArgLeuThrAsnlyValGluPheProPheLeuVal 20
DB 672 ATGAGAGCTCATGCACTTACTATTAGTGACCAATAAGTAGAATTTCTTTTACTT 731
QY 21 LeuLeuIleserGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
DB 732 CTTTGAATTTCTGAGGCTGCTGCTTTGGCATTAGTCAAGAGTTTCAAGATTTCTG 791
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
DB 792 CTTCTTGGAAGCTTTTGACATAGCACAGTCACTGTTGACAGAGTGGCAAGAGA 851
QY 61 LeuSerLeuIleLysHisProGlyCysSerThrMetSerGlyGlyLysAlaIleGlnHis 80
DB 852 CTTTCTTTAATAAAAACATCCAGAGTGTCCACCATGAGTGGGGAAGCCATAGACAT 911
QY 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLys 100
DB 912 TTGGCCCAACAAAGAAATAGATTTCATTTTGCATCAATCAAACTCCCTTCATCATCTAAA 971
QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
DB 972 AATTGTATTTTCTTTTACTGACTTCAACACGTTACTGATAAATATATATGAAAAAG 1031
QY 121 GluLysGlnGlyGlyIleGlyLysGlyClnIleLeuSerSerAlaAlaAspIleAlaAla 140
DB 1032 GAAAGAAAGAGAGATTTGAGAGAGGCAAAATCTGTCTTTCAGCAGCAGACATCTGCC 1091
QY 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
```

```
DB 1092 ACAGTACGACACACATGGCATGTCATCTTGTGAAAAACACATCGGCTATTTCTGTT 1151
QY 161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyVal 180
DB 1152 TGTAAAGCAGAGAGACTTGTACTTCAAAATATATGAGTACTGTTGATCTGTGTGTGC 1211
QY 181 AlaSerAspPheThrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr 200
DB 1212 GCAAGTAACTTCTATATCCGAGAGCTCTGGAAATTTTAAACAAACGCAACACAGTCACT 1271
QY 201 LeuLeuCysProProProArgLeuCysThrAspAsnGlyIleMetIleAlaThrAsnGly 220
DB 1272 TTGTGTGCTCTCTCCCAAGCTATGACATGATTAATGCAATGATTCATGAGATGCT 1331
QY 221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGlyGlyIleArgTyArg 240
DB 1332 ATTGAAAGACTACGCTGCTGGCTTGGGCAATTTTACATGACATAGAGGCAATCGCTATGA 1391
QY 241 ProlyscysProleuglyValaspilserlysglyValglyValasertilelyVal 260
DB 1392 CCAAAATGCTCTTGAGAGTACATATCAAAAGAGTTGAGAGAGCTTCCATAAAGTA 1451
QY 261 ProglinleuylMetglulle 267
DB 1452 CCACAATTAAAAATGAGATA 1472

RESULT 4
US-10-067-443-21
; Sequence 21, Application US/10067443
; Patent No. 6642041
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE HIGHLY EXPRESSED IN
; TITLE OF INVENTION: SPINAL CORD, MP-1
; FILE REFERENCE: D0073 NP
; CURRENT APPLICATION NUMBER: US/10/067,443
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 21
; LENGTH: 1387
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-067-443-21

Alignment Scores:
Pred. No.: 9,58e-174 Length: 1387
Score: 1358.50 Matches: 266
Percent Similarity: 91.10% Conservative: 0
Best Local Similarity: 91.10% Mismatches: 1
Query Match: 98.09% Indels: 25
DB: 4 Gaps: 1

US-10-649-273-2_COPY_148_414 (1-267) x US-10-067-443-21 (1-1387)
QY 1 MetglualahisAlaLeuThrIleArgLeuThrAsnlyValGluPheProPheLeuVal 20
DB 465 ATGAGAGCTCATGCACTTACTATTAGTGACCAATAAGTAGAATTTCTTTTACTT 524
QY 21 LeuLeuIleserGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
DB 525 CTTTGAATTTCTGAGGCTGCTGCTTTGGCATTAGTCAAGAGTTTCAAGATTTCTG 584
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
DB 585 CTTCTTGGAAGCTTTTGACATAGCACAGTCACTGTTGACAGAGTGGCAAGAGA 644
QY 61 LeuSerLeuIleLysHisProGlyCysSerThrMetSerGlyGlyLysAlaIleGlnHis 80
```

|||||
Db 645 CTTTCTTATATAAATCCAGAGTCTCCACCATGAGTGTGGAAAGCCATAGACAT 704
OY 81 LeuAlaLySGInGlyAsnArgPheHisPheAspIleYsProPheUhiSHiSAlaLys 100
Db 705 TTGGCCAAACAGAAATAGATTTCATTTCATTCATCAACCTCCCTTGCACTCATGCTAAA 764
OY 101 AsnCyAspPheSerPheThrgIleuGInHisValThraPlySleIleMetLysLys 120
Db 765 AATTGATTTTCTTTTACTGACCTCAACGCTTACTATAAATATATGAAAAAG 824
OY 121 GluLySGInGlyLysIle----- 126
Db 825 GAAAAAGAGAAAGTATATTCTAATTAGTAAAGATTAAACAGATTAATTCCTGATG 884
OY 127 -----GluLySGInGlyLysSerSerAla 135
Db 885 TGCCATAAAATAGCTGCTCATTTCTGCAAGTATGAGAAAGGCAATCCTGTCTTCAGCA 944
OY 136 AlaAspIleAlaAlaThraValGInHisThraMetAlaCySHiSLeuValLysArgThriS 155
Db 945 GAGACATTGCTGCCACAGTACAGCACATGSCATGTCATCTTGGAAGAAAGACAT 1004
OY 156 ArgAlaIleLeuPheCySlySGInArgAspLeuProGInHisAsnAlaValLeuVal 175
Db 1005 CGGGCTATTCTGTTTGTAAAGCAGAGACTGTTTACTCAAAATATATGACATCTGTT 1064
OY 176 AlaSerGlyGlyValAlaSerAsnPheTrIleArgArgAlaLeuGInIleLeuThraSrn 195
Db 1065 GCAATCGGTGGTGTCCAAAGTAACTTCTATATCCGAGACTCTGGAATTTTAAACAAC 1124
OY 196 AlaThrgInCyThraPheLeuCySProProArgLeuCySThraPAsnGlyIleMet 215
Db 1125 GCAACACAGTGCCTTTGTTGTCTCTCCCTCCAGACTATGCACTGTAATGGCATTAATG 1184
OY 216 IleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeuGInIleLeuHisAspIleGlu 235
Db 1185 ATTCAGTGAATGTATGAAAGCTACGTGTGGCTTGGCAATTTACATGACATAGAA 1244
OY 236 GlyIleArgTrGluProLysCySProLeuGlyValAlaSerIleSerGlyValGlyGlu 255
Db 1245 GGCATCCGTATGAAACCAAAATGCTCTTGAGTAGACATATCAAAAGAGTTGAGAA 1304
OY 256 AlaSerIleLysValProGInLeuLysMetGluIle 267
Db 1305 GCTTCATATAAGTACCAATTAATAATGAGATA 1340
RESULT 5
US-10-067-443-20
; Sequence 20, Application US/10067443
; Patent No. 6642041
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE HIGHLY EXPRESSED IN
; FILE OF INVENTION: SPINAL CORD, MP-1
; FILE REFERENCE: D0073 NP
; CURRENT APPLICATION NUMBER: US/10/067,443
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 14364
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-067-443-20
Alignment Scores:
Pred. No.: 1,34e-99 Length: 14364
Score: 824.50 Matches: 240

Percent Similarity: 26.73% Conservative: 0
Best Local Similarity: 26.73% Mismatches: 2
Query Match: 59.53% Indels: 657
DB: 4 Gaps: 4
US-10-649-273-2_copy_148_414 (1-267) x US-10-067-443-20 (1-14364)
OY 1 MetGluLahisAlaLeuThriIleArgLeuThraLysValGluPheProPheLeuVal 20
Db 10845 ATGAGGCTCATGCACTTACTATTAGGTGACCAATTAAGTAGAATTTCTTTTGTGTT 10904
OY 21 LeuLeuIleSerGlyGlyHisCySLeuLeuAlaLeuValGInGlyValIserAspPheLeu 40
Db 10905 CTTTGAATTTCTGAGGTCTACCTGTGTGGCATTAAGTCAAGAGTTTCAGATTTCTG 10964
OY 41 LeuLeuGlyLysSerLeuAspIleAlaProGInYasPheLysAspLysVal----- 57
Db 10965 CTTATGGAAGTCTTTGACATATGACACAGGTACATGCTTGACAAAGT-AAATTAAGAA 11023
OY 57 ----- 57
Db 11024 TTATTTCTGCATTTCTTTTGTATGTGTTCATTTCACTTAAGTAGCAATGATGTC 11083
OY 57 ----- 57
Db 11084 TACCACATTCACTTAATATTTCTGAATTTTATCTTAGTAACTGAAAAAATTCACAT 11143
OY 57 ----- 57
Db 11144 ATGTGAGAAAAAATAGAAAGTAGTAGACAAATTTATTAATCTTAGCTTTCTTAATA 11203
OY 57 ----- 57
Db 11204 AAATGTAAAGATTCTATCTATCTGTACATAAAGGCTGAATAAGTTTGACATACATTATG 11263
OY 57 ----- 57
Db 11264 TATTTGCCAAATATATGTATGTAAAGAAAGTGTCTGTAACTAATACATCTGCAAAAAA 11323
OY 57 ----- 57
Db 11324 GGTAAATTAAGAAAT 11383
OY 57 ----- 57
Db 11384 TTAATACACAAATTACTTACACACAGAGGTCCTCCCTCCCTCTTTGTTTAG 11443
OY 57 ----- 57
Db 11444 AATATACAGAGCTACTGCATATATAGAAAAAACAACAACAACAACAACAACAACA 11503
OY 57 ----- 57
Db 11504 CTTCCACAGTGAAT 11563
OY 57 ----- 57
Db 11564 AAGAGTTATTTCACTTCAAGCCATTTTCCAAACATATGAAGCAAAATATAGACAG 11623
OY 57 ----- 57
Db 11624 GGGGAGTATGGCTCTTATATTTGCGGTATCATATAAGAAACAGGTTGTCTGTAC 11683
OY 57 ----- 57
Db 11684 TGAATATCAGTAT 11743
OY 57 ----- 57
Db 11744 TTTTGTGTAT 11803
OY 58 -----AlaArgLeuSerLeuIle 64
|||||

CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 884
LENGTH: 1059
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-884

Alignment Scores:
Pred. No.: 3,67e-29 Length: 1059
Score: 291.50 Matches: 87
Percent Similarity: 50.41% Conservative: 37
Best Local Similarity: 35.37% Mismatches: 99
Query Match: 21.05% Indels: 23
Gaps: 9

US-10-649-273-2_COPY_148_414 (1-267) x US-09-252-991A-884 (1-1059)

1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLys---ValGluPheProPheLeu 19
711 ATGGAAGGCGACCTGCGCGCCGATGCTGGAAGACGACCGCGCGTCCCGTTCGTC 652
20 ValLeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPhe 39
651 GCCTTGCTGTTTCCGGGTCACACCCAGTGTGGGGTGAAGGATATCGGCGCTAC 592
40 LeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 59
591 CAGTTGCTTGGCAATCGGTGACGATGCCGCGCGCAACCTTCGACAAAGCCGCAAG 532
60 ArgLeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyValAlaIleGlu 79
531 CTGATCGGCGCT---GGCTATCCC-----GGTGTCCGGAATCGCC 493
80 HisLeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProPheLeuHisAla 99
492 CCCTGCGCGAGCGCGGACCTCTGCGCGCTGTGTTCCCGCGCGGATGACCGATCGC 433
100 LysAsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLys 119
432 CCGCGCTGACCTTCAAGCGGCTTCAAGACCTTCAAGCTTCAAGCTTCAAGCTTCAAG 376
120 LysGluLysGluGluGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 139
375 CGTTCGCTGAGGCGCGCGGACGACGACGACGACGACGACGACGACGACGACGACGAC 325
140 AlaThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeu 159
324 CTGGCGCTTCACAGCGCGGTGTGAGACCTCTGATCAAGTCCGCTCGCGCTTGC--- 268
160 PheCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGly 179
267 -----AAGCAGACCGGCGCT---AAGAAC-----CTGATGATCGCGGCGGT 229
180 ValAlaSerAspPheThrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCys 199
228 GTCAAGCGCAACGAGCGCTGCGGCTGGAAGATGCTCGCGCGGAATGAAGGG 169
200 ThrLeuLeuCysProProPheArgLeuCysThrAspAsnGlyIleMetIleAlaTrpAsn 219
168 CAGGTTCCTACGCGCGCGCGCTTCTGACGACGATGCGGATGATGCGCTACGCC 109
220 GlyIleGlnArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyr 239
108 GGTTCGCGCGCTGCGCGCGC-----CAGCATGACGCGCGCGGATGACGCGTGC 58
240 GluProLysCysProLeu 245
57 CAGCGCGCTGCGCGATG 40

RESULT 9
US-09-252-991A-801
Sequence 801, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 801
LENGTH: 1206
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-801

Alignment Scores:
Pred. No.: 4.55e-29 Length: 1206
Score: 291.50 Matches: 87
Percent Similarity: 50.41% Conservative: 37
Best Local Similarity: 35.37% Mismatches: 99
Query Match: 21.05% Indels: 23
Gaps: 9

US-10-649-273-2_COPY_148_414 (1-267) x US-09-252-991A-801 (1-1206)

1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLys---ValGluPheProPheLeu 19
514 ATGGAAGGCGACCTGCGCGCCGATGCTGGAAGACGACCGCGCGTCCCGTTCGTC 573
20 ValLeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPhe 39
574 GCCTTGCTGTTTCCGGGTCACACCCAGTGTGGGGTGAAGGATATCGGCGCTAC 633
40 LeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 59
634 CAGTTGCTTGGCAATCGGTGACGATGCCGCGCGCAACCTTCGACAAAGCCGCAAG 693
60 ArgLeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyValAlaIleGlu 79
694 CTGATCGGCGCT---GGCTATCCC-----GGTGTCCGGAATCGCC 732
80 HisLeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProPheLeuHisAla 99
733 CGCTGCGCGAGCGCGGACCTCTGCGCGCTGTGTTCCCGCGCGGATGACCGATGCC 792
100 LysAsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLys 119
793 CCGCGCTGACCTTCAAGCGGCTTCAAGACCTTCAAGCTTCAAGCTTCAAGCTTCAAG 849
120 LysGluLysGluGluGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 139
850 GCTTCGCGCGAGCGCGGACGACGACGACGACGACGACGACGACGACGACGACGAC 900
140 AlaThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeu 159
901 CTGGCGCTTCACAGCGCGGTGTGAGACCTCTGATCAAGTCCGCTCGCGCGTTC--- 957
160 PheCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGly 179
958 -----AAGCAGACCGGCGCT---AAGAAC-----CTGATGATCGCGGCGGT 996
180 ValAlaSerAspPheThrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCys 199
997 GTCAAGCGCAACGAGCGCTGCGGCTGGAAGATGCTCGCGCGGAATGAAGGG 1056

Oy	200	ThrlauLeaucysPrcProArleuThrlaArgleucysTrAspAsglylleMecllaeATPaSn	219
Db	1057	CAGGTTCTACGCCCGCCGGCCTTCTCACCGCAATGCGCGCATGTGCCTTAAGCC	1116
Oy	220	GlylllegluArgleuArgllaAglyleuglylleuehlAspilleguglylleArgTyR	239
Db	1117	GGTCGCGAGCGGCTGTGCGCGGC-----CACGAAGAAGCGCGCGCATCAGGTC	1167
Oy	240	GlUpPolycysProlen	245
Db	1168	CAGCGCGCTGCGCGCATG	1185
RESULT 10			
US-09-557-884-1/c			
Sequence 1, Application US/09557884			
Patent No. 6506581			
GENERAL INFORMATION:			
APPLICANT: Fleischmann et al.			
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof			
NUMBER OF SEQUENCES: 1			
CORRESPONDENCE ADDRESS: ADDRESSEE: Human Genome Sciences, Inc. STREET: 9410 Key West Avenue CITY: Rockville STATE: MD COUNTRY: USA ZIP: 20850			
COMPUTER READABLE FORM: MEDIUM TYPE: 3 1/2 inch diskette COMPUTER: Dell Pentium OPERATING SYSTEM: MS DOS V6.22 SOFTWARE: ASCII Text			
CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/557,884 FILING DATE: 25-Apr-2000 CLASSIFICATION: <Unknown>			
PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/476,102 FILING DATE: JUN-5-1995			
ATTORNEY/AGENT INFORMATION: NAME: Michelle S. Marks REGISTRATION NUMBER: 41,971 REFERENCE/DOCKET NUMBER: PB186P3 TELECOMMUNICATION INFORMATION: TELEPHONE: 301-309-8504 TELEFAX: 301-309-8439			
INFORMATION FOR SEQ ID NO: 1: SEQUENCE CHARACTERISTICS: LENGTH: 1830121 base pairs TYPE: nucleic acid STRANDEDNESS: double TOPOLOGY: linear			
SEQUENCE DESCRIPTION: SEQ ID NO: 1: US-09-557-884-1			
Alignment Scores:			
Pred. No.:	1,68e-23	Length:	1830121
Score:	289.50	Matches:	82
Percent Similarity:	48.70%	Conservative:	30
Best Local Similarity:	35.65%	Mismatches:	95
Query Match:	20.90%	Indels:	23
DB:	4	Gaps:	7
US-10-649-273-2_COPY_148_414 (1-267) x US-09-557-884-1 (1-1830121)			
Oy	1	MecgluaLahIsAlaleuThrlleaArgleu---ThrAsnLysValGlUpheProPhelu	19
Db	552881	ATGGAAGGCAATTACTTGCCGCAATGCTTGATGACAATTCACCGCAGCATTTCTTTGTT	552822
Oy	20	VallouLenuLseGlyGlyhisCySeuleuAlaleuValGlnGlyValSeraspHe	39

```

Db      552821  GCTCGTTGGATCGGGTGGCCACACTCAATTAGTGCNGTGCATGCTGAGGAAATAT 552762
Oy      40      LeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 59
           :::::|||||
Db      552761  GAAAGTAGGAGAAATCTTAATGATATGCTGTGGCGAAGCCTTTCATTAACAGAGAAA 552702
Oy      60      ArgLeuSerLeuIleLeuHisProGlyCysSerThrMetSerGlyGlyLysAlaIleGlu 79
           |||||::|
Db      552701  TTACTTGGACATCA--GATTATCA-----GGTGGCCGGCAGCTTTCT 552653
Oy      80      HisLeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAla 99
           |||||::|
Db      552662  CGTTTAGGGGAAAAAGTAAACCGCAATCGTTTCACATTTCCACGTCGAATGACAGATCGT 552603
Oy      100     LysAsnCyAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLys 119
           :::::|||||
Db      552602  GCAGCGCTTGATTTTATGATTTTCTCGTTTAAACATTTGCCGCAATACAGTTAATCA 552543
Oy      120     LysGlyLysGlnGluGly-----IleGlyLysGlnIleLeuSerSerAlaAlaAsp 137
           |||||::|
Db      552542  GCATTTAAACAGAGGGCGCAACTGATAGAGCA-----ACTAAAGCAGAT 552498
Oy      138     IleAlaIaIaThrValGlnHisIleThrMetAlaCysHisLeuValLysArgThrHisArgAla 157
           |||||::|
Db      552497  ATTGCTTATGCTTTCCAAAGATCGGTGGTGCATACCTTGGC----- 552456
Oy      158     IleIeuPheCysLysGlnArgAspLeuProGlnAsnAsnAlaValLeuValAlaSer 177
           :::::|||||
Db      552455  --ATTAAATGTAAAG--CGTGCAATTAAGAAAGAAACAGCCTTAATAACGTTAAGATTGGC 552402
Oy      178     GlyGlyValAlaSerAsnPheTyrIleArgAlaLeuGluIleLeuThrAsnAlaThr 197
           |||||::|
Db      552401  GGAGGGGTGAGGCGCAATTAACCTCCGAGAAAGCCTTGGCCCACTTAATGCAAAATTTA 552342
Oy      198     GlnCysThrLeuLeuCysProProArgLeuCysThrAspAsnGlyIleMetIleAla 217
           :::::|||||
Db      552341  GGTGGCGGAAGTCTTATTCCTCAACTCAATTTGTACAGATTAATGTCGATGATTGCT 552282
Oy      218     TrpAsnGlyIleGluArgLeuArgAlaGly 227
           :::::|||||
Db      552281  TACACAGGTTTTTACGTTTAAACAGGT 552252

RESULT 11
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
GENERAL INFORMATION:
APPLICANT: Robert D. Fleischmann
           Mark D. Adams
           Owen White
           Hamilton O. Smith
           J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of
                    the Haemophilus influenzae Rd Genome, Fragments
                    Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Alignment Scores:
Pred. No.: 1,68e-23 Length: 1830121
Score: 289.50 Matches: 82
Percent Similarity: 48.70% Conservative: 30
Best Local Similarity: 35.65% Mismatches: 95
Query Match: 20.90% Indels: 23
DB: 4 Gaps: 7

US-10-649-273-2_COPY_148_414 (1-267) x US-09-643-990A-1 (1-1830121)
QY 1 MetGUAhAhIsAlaLeuThrIleArgLeu---ThrsAluValGluPheProPheLeu 19
DB 552881 ATGGAGGCGCATTTACTGGCCAAATCGTTGAGACATTCACCGCATTTCTTTGTT 552822
QY 20 ValLeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPhe 39
DB 552821 GCTCTGTGGTATCGGGTGGCCACATCATTAAGTGGTGCATGAGTGAAGAAATAT 552762
QY 40 LeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspValAlaArg 59
DB 552761 GAAGTATAGAGAAATCTATGATGATGCTGCGCCAAACCTTTGATAAACGCAAAA 552702
QY 60 ArgLeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyValAlaIleGlu 79
DB 552701 TTACTTGACTA---GATTATCCA-----GGTGGCGGGCATTTCT 552663
QY 80 HisLeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProPheLeuHisAla 99
DB 552662 CGTTTACGGAAAGGTAGCCCAATCGTTTCACATTCACAGTCCAAATGACAGATCGT 552603
QY 100 LysAsnCysAspSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLys 119
DB 552602 GCAAGCGCTGATTAGTTTTCGTTTAAACATTTCCCGCAATACAGTTAAATCAA 552543
QY 120 LysGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLys 137
DB 552542 GCAATTAATAAACAAGGCGCACTAGTAGCAAA-----ACTAAAGCAGAT 552498
QY 138 IleAlaIleThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisAla 157
DB 552497 ATTGCTTATGCTTCCAAAGATGCGGTGGATGCTTTCG---552456
QY 158 IleLeuPheCysLysGlnArgAspLeuProGlnAsnAlaValLeuValAlaSer 177
DB 552455 ---ATTAAATGTAAG---CGTGCATTGAAGAAACAGGCTATAAAGCTTATGATGGCG 552402
QY 178 GlyGlyValAlaSerAsnPheTyrIleArgArgAlaLeuGluIleLeuThrAsnAlaThr 197
DB 552401 GGAGGGGTGAGCCCAATTAATAAATCCGAGAAACGCTTGGCAGCTTAATGCAAAATTTA 552342
QY 198 GlnCysThrLeuLeuCysProProArgLeuLeuThrAspAsnGlyIleMetIleAla 217
DB 552341 GGTGGCGAAGTGTATTATCTCAACCTCAATTTTGTACAGATTAAAGGCGAGATATGCT 552282

QY 218 TrpAsnGlyIleGluArgLeuArgAlaGly 227
DB 552281 TACACAGTGTTTTTCACCTTAAACAAGGT 552252

RESULT 12
US-09-902-540-6612
Sequence 6612, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Miegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 6612
LENGTH: 996
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-6612

Alignment Scores:
Pred. No.: 6.44e-28 Length: 996
Score: 282.00 Matches: 86
Percent Similarity: 48.93% Conservative: 28
Best Local Similarity: 36.91% Mismatches: 87
Query Match: 20.36% Indels: 32
DB: 4 Gaps: 9

US-10-649-273-2_COPY_148_414 (1-267) x US-09-902-540-6612 (1-996)
QY 1 MetGUAhAhIsAlaLeuThrIleArgLeuThrAsnVal---GluPheProPheLeu 19
DB 334 CTGAGAGGCGCATCTGCTGGCCATCGGCTTTGAGAGTGGCGCGGACCCCGCTTCTT 393
QY 20 ValLeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPhe 39
DB 394 GGGCTGCTGCTTCCGGGGGACACACAGCCTTACAGAGTGGCAGGCTTACGGGAGTAC 453
QY 40 LeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspValAlaArg 59
DB 454 CGGCTGGGGGACACACCGCGGACGCGCGCGGCGGAGGATGACAGACCGCTCGC 513
QY 60 ArgLeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyValAlaIleGlu 79
DB 514 ATCTCGGCGCTG---CCGATATCC-----GGTGGCGAGCCCATGCAC 552
QY 80 HisLeuAlaLysGlnGlyAsn-----ArgPheHisPheAspIleLysProPro 95
DB 553 CAGTTGGCGGACAGGGAACCCGAGGCCATCGGCTTC-----CGGGC 597
QY 96 LeuHisAlaLysAsnCysAspSerPheThrGlyLeuGlnHisValThrAspLys 115
DB 598 GCGGCGCGGCGCAACTTGCAGCTGCTTCCGGGTGAAG-----ACG 645
QY 116 IleIleMetLysLysGluLysGluLysGluLysGluLysGluLysGluLysGluLys 135
DB 646 GCGGCTGCGCAACGTCAGAGACCGCGCGCGGCGGAGGCGCGCTG-----696
QY 136 AlaAspIleAlaIleThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHis 155
DB 697 GCGGATTTGCGGCTCTTCCAGAGGCGGTGGGACGTCCTTCCAGAGAG-----750
QY 156 ArgAlaIleLeuPheCysLysGlnArgAspLeuProGlnAsnAlaValLeuVal 175
DB 751 -----CTGCGGCGCGCGCGCGCGGTG-----GCCACAGACGATGGTG 792

QY 176 AlaSerGlyValAlaSerAspPheThrIleArgAlaValLeuGluIleLeuThrAsn 195
DB 793 CTGTGGCGGGCGCTCTTCCAGAGCCGCTGCGAAGAG----- 779
QY 196 AlaThrGlnCysThrLeuLeuGluCysProProProArgLeuGluCysThrAspAsnGlyIleMet 215
DB 853 GAGCGGGGGGTGAACATGTTCTGCCCCCGGTGCGGTGCAAGCAATGGCGCATG 912
QY 216 IleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeu 228
DB 913 ATTGGCGTGGCGGGGTATGAGGGCTACCGCGGGCGCTG 951
RESULT 13
US-09-902-540-503/c
Sequence 503, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 503
LENGTH: 2582
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-503

Alignment Scores:
Pred. No.: 3,156-27 Length: 2582
Score: 282.00 Matches: 86
Percent Similarity: 48.93% Conservative: 28
Best Local Similarity: 36.91% Mismatches: 87
Query Match: 20.36% Indels: 32
DB: 4 Gaps: 9

US-10-649-273-2_COPY_148_414 (1-267) x US-09-902-540-503 (1-2582)

QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnIysVal---GluPheProPheLeu 19
DB 1195 CTGGAGGGGCGCTGCTGGCCATCGGCTGTGAGGTGGCGCGGAGCGCGCTTCTT 1136
QY 20 ValLeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPhe 39
DB 1135 GGGCTGTGCTTCCGGCGGGGACACAGGCTCTACAGAGTGCAGGCTTACGGGCACTAC 1076
QY 40 LeuLeuLeuGlyIysSerLeuAspIleAlaProGlyAspMetLeuAspIysValAlaArg 59
DB 1075 CGGCTGTGGGGCGACGCGCGACGACGCGCGCGGACGACATATGACAGAGCGCTCGC 1016
QY 60 ArgLeuSerLeuIleIysHisProGluCysSerThrMetSerGlyIysValAlaIleGlu 79
DB 1015 ATCTTCGGCTG---CGTATCCG-----GGTGGGAGCCCATGAC 977
QY 80 HisLeuAlaIysGlnGlyAsn-----ArgPheHisPheAspIleIysProPro 95
DB 976 CAGTTGGCGGAGAGGGGAAACCGGAGGCGCATCGCTTC-----CGCGC 932
QY 96 LeuHisHisAlaIysAsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspIys 115
DB 931 GGGCTCGCGGAGCAACTTCGACGCTCTCTCCGCGTTGAG-----ACG 884
QY 116 IleIleMetLeuIysGluGlyGlnGlyIleGluIysGlnIleLeuSerSerAla 135
DB 883 GCGGTGTCTCACACATGACAGAGCAAGCGGCGCGGCGGCGGCTG-----833
QY 136 AlaAspIleAlaIleThrValGlnHisThrMetAlaCysHisLeuValIysArgThrHis 155

DB 832 GCGGATTTGGCGGGCTCTTCCAGAGCCGCTGCGAAGAG----- 779
QY 156 ArgAlaIleLeuPheCysIysGlnArgAspLeuLeuProGlnAsnAlaValLeuVal 175
DB 778 -----CTGGGCGCGCGCGCGCGCTG-----GCCACAGCACTTG 737
QY 176 AlaSerGlyValAlaSerAspPheThrIleArgAlaValLeuGluIleLeuThrAsn 195
DB 736 CTGTGGCGGGCGCTGCGCGAAGCTCGGCTGCGGACATGTGTACAGCGCGAGCCGAG 677
QY 196 AlaThrGlnCysThrLeuLeuGluCysProProProArgLeuGluCysThrAspAsnGlyIleMet 215
DB 676 GAGCGGGGGGTGAACATGTTCTGCCCCCGGTGCGGTGCAAGCAATGGCGCATG 617
QY 216 IleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeu 228
DB 616 ATTGGCGTGGCGGGGTATGAGGGCTACCGCGGGCGCTG 578

RESULT 14
US-09-543-681A-2341
Sequence 2341, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETTON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709,1002-001
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 2341
LENGTH: 1074
TYPE: DNA
ORGANISM: Proteus mirabilis
US-09-543-681A-2341

Alignment Scores:
Pred. No.: 1,176-27 Length: 1074
Score: 280.50 Matches: 81
Percent Similarity: 46.77% Conservative: 35
Best Local Similarity: 32.66% Mismatches: 101
Query Match: 20.25% Indels: 31
DB: 4 Gaps: 8

US-10-649-273-2_COPY_148_414 (1-267) x US-09-543-681A-2341 (1-1074)

QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnIysVal---GluPheProPheLeu 19
DB 385 ATGAGAGGCACTTATTATGAGCGCGCATGCTGAAGAAAGAACCGCAATTTCTTCTG 444
QY 20 ValLeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPhe 39
DB 445 GCGTATTGTCTCGGGGGGACATACACAACTTATTGTGTAAACAGGATGGGGAATAT 504
QY 40 LeuLeuLeuGlyIysSerLeuAspIleAlaProGlyAspMetLeuAspIysValAlaArg 59
DB 505 ACCCTGTAGGTGATGCTATTGATGATGCTGCTGTAAGCATTTGTAATAACAGCCAG 564
QY 60 ArgLeuSerLeuIleIysHisProGluCysSerThrMetSerGlyIysValAlaIleGlu 79
DB 565 CTATTGGGGCTT---GATTATCC-----GGCGGCGCTGTTTATCA 603
QY 80 HisLeuAlaIysGlnGlyAsnArgPheHisPheAspIleIysProProLeuHisHisAla 99
DB 604 AATAATGCACACAGAGGTGAAGAGACGTTTCTTTCTGCTCCCATATGACAGACA 663
QY 100 IysAsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspIleIleMetIys 119
DB 664 CCGGACATCACTTATGTTCTGTGTTAAACCTTGGCCGCTATATATATGTCGCA 723

```

QY 120 LysGluLysGluGluGlyLeuGlyGlyGlnLeuSerSerAlaAlaSerIleAla 139
DB 724 AACGATGATTCAGG-----CAAACTCGAGCAGATATTGGC 759
QY 140 AlaThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeu 159
DB 760 CGTCTTTTGAAGATGCGGTAGATATCTTGGCAATAAATGTCGTGCGCA----- 813
QY 160 PheCysLysGlnArgAspLeuLeuProGlnAsnAsn--AlaValLeuValAspGly 178
DB 814 -----TTAGACCAACAGCCTTTAAACCTTAGTATGATGCTGG 852
QY 179 GlyValAlaSerAspPheTyrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGln 198
DB 853 GGGGTAGTACTTAACGCTACTTACGCGCAAAATGCGATGATAGAACAACTCGCA 912
QY 199 CysThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 218
DB 913 GGGGAAGTGTATTATGCTCGCCCTGAGTTATGTACCATATATGTCATATGCTTTG 972
QY 219 AsnGlyIleGluArgLeuArgAlaGlyLeu--GlyIleLeuHisAspIleGluGlyIle 237
DB 973 GCGGAGTGTATCCGTTTAAAGGTGATACGAGGCGCTTTA-----GGGGTG 1020
QY 238 ArgTyrGluProLysCysProLeu 245
DB 1021 ACAGTGAGACCACTGCGCTTTA 1044

RESULT 15
US-08-087-797-1
; Sequence 1, Application US/08087797
; Patent No. 5543312
GENERAL INFORMATION:
APPLICANT: Mellors, Alan
APPLICANT: Lo, Reggie Y. C.
APPLICANT: Abdullah, Khalid M.
TITLE OF INVENTION: Pasteurella Haemolytica
TITLE OF INVENTION: Glycoprotease
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSER: Bell, Seltzer, Park & Gibson, P.A.
STREET: 1211 East Morehead Street,
CITY: Charlotte
STATE: No. 5543312th Carolina
COUNTRY: United States
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,797
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Layton, Jr., Samuel G.
REGISTRATION NUMBER: 22807
REFERENCE/DOCKET NUMBER: 3374-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 704 377 1561
TELEFAX: 704 334 2014
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1315 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: 141..1115
US-08-087-797-1

Alignment Scores:
Pred. No.: 3.72e-26 Length: 1315
Score: 270.50 Matches: 79
Percent Similarity: 46.46% Conservative: 26
Best Local Similarity: 34.96% Mismatches: 102
Query Match: 19.53% Indels: 19
DB: 1 Gaps: 6

US-10-649-273-2_COPY_148_414 (1-267) x US-08-087-797-1 (1-1315)
QY 1 MetGluAlaHisAlaLeuThrIleArgLeu--ThrAsnLysValGluPheProPheLeu 19
DB 474 ATGGAAGGCGATTACTTACCTCCCAATGTTGGAAGAAAATGCCCGAATTCGTTTGG 533
QY 20 ValLeuLeuIleSerGlyIleHisCysLeuLeuAlaLeuValGlnGlyValSerAspPhe 39
DB 534 GCATTATTGATTTCAGTGCACACCCCACTGTTAAAGTTGACGCGCTTGGCGCAATAC 593
QY 40 LeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 59
DB 594 GAACCTACTCGGGGAATCAATTGATGATCTCGCGGTGAAGCCTTTGACAAACAGCGCAA 653
QY 60 ArgLeuSerLeuIleLysHisProGluCysSerThrMetSerGlyIleLysAlaIleGlu 79
DB 654 CTACTCGGTTTG--GATTACCT-----GCCGGTCTAGCATTGCA 692
QY 80 HisLeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisAla 99
DB 693 AAATTAGCGAATCGGACGCGCAAAATGTTTAAATTCCTCGTCAATGACGACAGA 752
QY 100 LysAsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLys 119
DB 753 CCGGACCTGGAATTCAGTTCTCCCGTTTAAAAACCTTTCCTGCAATACGATTAAAGCC 812
QY 120 LysGluLysGluGluGlyIleGluLysGlnIleLeuSerSerAlaAlaSerIleAla 139
DB 813 AATCTTAAATGAATGTGAATCGATGAGCAA-----ACCAATGCGATTATGCG 863
QY 140 AlaThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeu 159
DB 864 CACGATTCACCAACAGCCGCG-----GTTGATCTATTTTAATT 902
QY 160 PheCysLysGlnArgAspLeuLeuProGlnAsnAlaValLeuValAlaSerGlyIle 179
DB 903 AAATGCAAG--CGAGCGTTAGACCAACCGGCTATTAAACGCTTAGTATGCGACGCGC 959
QY 180 ValAlaSerAspPheTyrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCys 199
DB 960 GTAAGTGCCAATAAACAAATTAGACAGACCTTGGCGAATATGAAAAAATTAAAGGC 1019
QY 200 ThrLeuLeuCysProProArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 219
DB 1020 GAAGTATTCTACCTTCGCGCAATTTTGCATGACAGCGCGCAATGATTCCTACT 1079
QY 220 GlyIleGluArgLeuArg 225
DB 1080 GCGTTTCTTCGCTTAAA 1097

```

Search completed: June 17, 2005, 02:16:07
Job time : 1068.63 secs

GenCore version 5.1.6.
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 16, 2005, 23:18:53 ; Search time 1813.86 Seconds
(without alignments)
913.751 Million cell updates/sec

Title: US-10-649-273-2_COPY_148_414
Perfect score: 1385
Sequence: 1 MEAHALTRLTNKVEPFLV.....DISKEVGEASIKVQOLMEI 267

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6054689 segs, 3103772919 residues

Total number of hits satisfying chosen parameters: 12109378

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame.p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool.p/US10649273/runat.15062005.111418.6138/app.query.fasta.1.1429
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsu2
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODS=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10649273.OCGN.1.1.1034.0/runat.15062005.111418.6138
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEOUTRY -NEG SCORES=0 -WAIT -DSBLOK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FAPOP=6 -FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
- 3: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
- 4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
- 6: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
- 7: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
- 8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
- 10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
- 12: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
- 13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
- 14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
- 15: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
- 16: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
- 17: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
- 18: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
- 19: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
- 20: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
- 21: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
- 22: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
- 23: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
- 24: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
- 25: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
- 26: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1385	100.0	1416	US-10-120-988-177	Sequence 177, App
2	1385	100.0	1526	US-10-067-443-23	Sequence 23, App
3	1385	100.0	1526	US-10-649-273-23	Sequence 23, App
4	1385	100.0	1526	US-10-651-722-23	Sequence 23, App
5	1385	100.0	2197	US-10-067-443-1	Sequence 1, App
6	1385	100.0	2197	US-10-649-273-1	Sequence 1, App
7	1385	100.0	2197	US-10-651-722-1	Sequence 1, App
8	1385	100.0	2572	US-10-480-988-26	Sequence 36, App
9	1385	98.1	1387	US-10-067-443-21	Sequence 21, App
10	1385	98.1	1387	US-10-649-273-21	Sequence 21, App
11	1385	98.1	1387	US-10-651-722-21	Sequence 21, App
12	1385	98.1	1387	US-10-887-553A-1047	Sequence 1047, App
13	1348	97.3	1245	US-10-012-140-6	Sequence 6, App
14	1348	97.3	1820	US-10-012-140-6	Sequence 4, App
15	1204	86.9	2208	US-10-094-749-400	Sequence 400, App
16	1204	86.9	2890	US-10-723-860-7447	Sequence 7447, App
17	824.5	59.5	14364	US-10-067-443-20	Sequence 20, App
18	824.5	59.5	14364	US-10-649-273-20	Sequence 20, App
19	824.5	59.5	14364	US-10-651-722-20	Sequence 20, App
20	468	33.8	371	US-10-430-201-3118	Sequence 3118, App
21	468	33.8	371	US-10-430-201-3119	Sequence 3119, App
22	422.5	30.5	1917	US-10-424-599-66417	Sequence 66417, App
23	359	25.9	1628	US-10-437-963-11249	Sequence 11249, App
24	342	24.7	756	US-10-081-051-8	Sequence 8, App
25	342	24.7	4360	US-10-081-051-8	Sequence 2, App
26	321.5	23.2	1146	US-10-282-122A-14674	Sequence 14674, App
27	314.5	22.7	936	US-10-282-122A-8315	Sequence 8315, App
28	308.5	22.3	1044	US-10-282-122A-26972	Sequence 26972, App
29	308.5	22.3	94750	US-10-672-187-38	Sequence 38, App
30	304.5	22.0	1032	US-10-282-122A-31043	Sequence 31043, App
31	301.5	21.8	1000	US-10-343-561-50	Sequence 50, App
32	291.5	21.0	1026	US-09-815-242-7701	Sequence 7701, App
33	291.5	21.0	1026	US-10-282-122A-30016	Sequence 30016, App
34	289.5	20.9	1029	US-09-815-242-6946	Sequence 6946, App
35	289.5	20.9	1029	US-10-282-122A-22020	Sequence 22020, App
36	289.5	20.9	1830121	US-10-329-670-1	Sequence 1, App
37	289.5	20.9	1830121	US-10-158-865-1	Sequence 1, App
38	285	20.6	372	US-10-357-930-61088	Sequence 61088, App
39	280.5	20.3	1020	US-10-282-122A-32254	Sequence 32254, App
40	279.5	20.2	1014	US-09-815-242-9682	Sequence 9682, App
41	279.5	20.2	1014	US-10-282-122A-39301	Sequence 39301, App
42	278.5	20.2	1014	US-10-282-122A-41977	Sequence 41977, App
43	277.5	20.0	1007	US-10-282-122A-19320	Sequence 19320, App
44	277	20.0	927	US-10-282-122A-25354	Sequence 25354, App
45	277	20.0	1023	US-10-282-122A-31809	Sequence 31809, App

ALIGNMENTS

RESULT 1
US-10-120-988-177
; Sequence 177, Application US/10120988
; Publication No. US20030219745A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenchua
; APPLICANT: Ren, Feiyun
; APPLICANT: Wang, Dumei
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11

```
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pc_fl_genes Version 2.0
; SEQ ID NO 177
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (205)..(1305)
US-10-120-988-177

Alignment Scores:
Pred. No.: 1,18e-169 Length: 1416
Score: 1385.00 Matches: 267
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x US-10-120-988-177 (1-1416)

QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
DB ATGGAGGCTCATGCACTTACTATTAGGTGACCAATAGAGTAATTCCTTTTAGTT 561
QY 21 LeuLeuIleSerGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
DB CTTTATATTTCTGGAGCTCATCTGTCTGTGGCATTTAGTTCAAGAGTTTCGATTTCGTG 621
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
DB CTTCTTGGAAAGCTTTGGACATAGCACACGGTGACATGCTTGCAAGTGCGCAAGAGA 681
QY 61 LeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyValAlaIleGlnHis 80
DB CTTTCTTTATAATAAATCATCCAGAGTGTCTCCACCATGATGGTGGAAGCCATGAAACAT 741
QY 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProPheLeuHisAlaLys 100
DB TTGGCCAAACCAAGGAATATGATTTGATTTGACATCAACCTCCCTTGCAATGCTAA 801
QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValAlaThrAspLysIleIleMetLys 120
DB AATTGATTTTCTTTTACTGACCTTCAACACGTTACTGTAAATTAATATATATAAAG 861
QY 121 GluLysGlnGlnGlyIleGlnLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 140
DB GAAAGAGGAAAGGATTTGAGAAAGGGCAATCTGTCTTCAGACAGACATTTGCTGCC 921
QY 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
DB ACAGTACAGACACAATGCGATGTCATCTTGTGAAGAACACATCGGGCTAATCTGTTT 981
QY 161 CysLysGlnArgAspLeuLeuProGlnAsnAspAlaValLeuValAlaSerGlyVal 180
DB TGTAAACAGAGAACTTTGTAACCTCAAAATATGCAATGATCGTGTGATCGTGTC 1041
QY 181 AlaSerAspPheThrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr 200
DB GCAAGTACCTTCATATCCGACAGACTCTGGAATTTTAAACAAACGCAACGACATGCACT 1101
QY 201 LeuLeuCysProPheProArgLeuCysThrAspAsnGlyIleMetIleAlaTTPAsnGly 220
DB TTGTTGTGTCCTCTCCACAGACTATGACATGATATGACATTTAGTTGATGGAATGCT 1161
QY 221 IleGluArgLeuArgAlaGlyLeuGlyValLeuHisAspIleGlnGlyIleArgTyrGlu 240
DB ATTGAAGAGCTACGTGTGCTTGCGCATTTTACATGACATAGAGGACATCCGCTAAGAA 1221
QY 241 ProLysCysProLeuGlyValAspIleSerLysGlnValGlyValAlaSerGlyVal 260
DB |||||
```

```
DB 1222 CCAAAATGCTCTTGGAGTAGACATATCAAAAGAGTTGGAGAGCTTCATATAAAGTA 1281
QY 261 ProGlnLeuLysMetGluIle 267
DB |||||
DB 1282 CCACATTTAAATAATGAGATTA 1302

RESULT 2
US-10-067-443-23
; Sequence 23, Application US/10067443
; Publication No. US20030082782A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE HIGHLY EXPRESSED IN
; FILE REFERENCE: D0073 NP
; CURRENT APPLICATION NUMBER: US/10/067,443
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 1526
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-067-443-23

Alignment Scores:
Pred. No.: 1.32e-169 Length: 1526
Score: 1385.00 Matches: 267
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x US-10-067-443-23 (1-1526)

QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
DB 1 ATGGAGGCTCATGCACTTACTATTAGGTGACCAATAGAGTAATTCCTTTTAGTT 60
QY 21 LeuLeuIleSerGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
DB CTTTATATTTCTGGAGCTCATCTGTCTGTGGCATTTAGTTCAAGAGTTTCGATTTCGTG 120
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
DB CTTCTTGGAAAGCTTTGGACATAGCACACGGTGACATGCTTGCAAGTGCGCAAGAGA 180
QY 121 GluLysGlnGlnGlyIleGlnLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 140
DB CTTTCTTTATAATAAATCATCCAGAGTGTCTCCACCATGATGGTGGAAGCCATGAAACAT 240
QY 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
DB CTTTCTTTATAATAAATCATCCAGAGTGTCTCCACCATGATGGTGGAAGCCATGAAACAT 240
QY 161 CysLysGlnArgAspLeuLeuProGlnAsnAspAlaValLeuValAlaSerGlyVal 180
DB |||||
```


Db 481 TGTAAAGAGAGAGCTTTGTAACCTCAAAATATGACGTACTGTTGATCTGGTGTCTC 540
Qy 181 AAsenAspPheTyrIleArgArgAlaLeuGluIleuThraAsnAlaThrGlnCyThr 200
Db 541 GCAAGTAACCTTCTATATCCGACAGCTCTGGAAATTTTAAACAAACGACAGCTGACT 600
Qy 201 LeuLeuCyProProProArgLeuCyThrAspAsnGlyIleMetIleAlaTTPanGly 220
Db 601 TTGTTGTCTCTCTCCAGACTATGCACTGATTAATGCAATTATGATTCATGGAATGCT 660
Qy 221 TleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGlu 240
Db 661 ATTTGAAAGACTAGCTGCTGGCTTGGCAATTTTACATGACATGAGAGGCAATCCGTATGAA 720
Qy 241 ProLysCyProLeuGlyValAspIleSerIleGluValGlyGluAlaSerIleLeuVal 260
Db 721 CCAAAATGCTCTTGGAGTAAGACATATCAAAAGAACTGGAGAGCTTCCATTAAGTA 780
Qy 261 ProGlnLeuLysMetGluIle 267
Db 781 CCACAAATTAATAATGGAGATA 801
RESULT 3
US-10-649-273-23
; Sequence 23, Application US/10649273
; Publication No. US20040043407A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE, MP-1
; FILE REFERENCE: D0073 CNT
; CURRENT APPLICATION NUMBER: US/10/649,273
; PRIOR FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 10/067,443
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 1526
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-649-273-23
Alignment Scores:
Pred. No.: 1,32e-169 Length: 1526
Score: 1385.00 Matches: 267
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0
US-10-649-273-2_COPY_148_414 (1-267) x US-10-649-273-23 (1-1526)
Qy 1 MetGluAlaHisAlaLeuThrIleArgLeuThraAsnLysValGluPheProPheLeuVal 20
Db 1 ATGAGAGCTCATGCACTTACTATTAGGTTGACCAATAAGTAGAATTTCTTTTATTAGTT 60
Qy 21 LeuLeuIleSerGlyGlyHisCySerLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
Db 61 CTTTGAATTTCTGAGAGTCACTGCTGTGTGGCATTAATTCAGAGATTTTCAGATTTTCTG 120
Qy 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
Db 121 CTTCTTGAAAGGCTTTTGACATAGCACAGGTGACATGCTTGACAAAGGTGGCAAGAGA 180
Qy 61 LeuSerLeuIleLysHisProGluCySerThrMetSerGlyGlyLysAlaIleGluHis 80
Db 181 CTTTCTTTATTAATAACATCCAGAGTGTCTCCACCATGATGTGTGGAAAGGCATGAACAT 240
Qy 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisAlaLys 100

Db 241 TTGGCCAAACAGAGAAATAGATTTCATTTTGAATCAAACTCCCTTCATCATGCTAA 300
Qy 101 AsnCyAspPheSerPheThrGlyLeuGlnHisValThraAspLysIleIleMetLysLys 120
Db 301 AATTGTGATTTTCTTTTACTGACCTTCACACGCTACTGATTAATAATATGAAAAAG 360
Qy 121 GluLysGluGluGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaAla 140
Db 361 GAAAAAGAGAAAGGATATTGAAAGAGGCAAACTGCTCTTACAGCAGCAGACATGCTGCC 420
Qy 141 ThrValGlnHisThrMetAlaCyHisIleuValLysArgThrHisArgAlaIleuPhe 160
Db 421 ACAGTACAGACACAAATGGCAATGTCATCTTGTGAAAAGAACACATGGGCTATTTCTGTT 480
Qy 161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuAlaIleSerGlyGlyVal 180
Db 481 TGTAAAGAGAGAGCTTTTACTTCCATAATATATGACATGCTGTTGATCTGGTGTCTC 540
Qy 181 AAsenAspPheTyrIleArgArgAlaLeuGluIleuThraAsnAlaThrGlnCyThr 200
Db 541 GCAAGTAACCTTCTATATCCGACAGCTCTGGAATTTTAAACAAACGACACAGTGCAT 600
Qy 201 LeuLeuCyProProProArgLeuCyThrAspAsnGlyIleMetIleAlaTTPanGly 220
Db 601 TTGTTGTCTCTCTCCAGACTATGCACTGATTAATGCAATTATGATTCATGGAATGCT 660
Qy 221 TleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGlu 240
Db 661 ATTTGAAAGACTAGCTGCTGGCTTGGCAATTTTACATGACATGAGAGGCAATCCGTATGAA 720
Qy 241 ProLysCyProLeuGlyValAspIleSerIleGluValGlyGluAlaSerIleLeuVal 260
Db 721 CCAAAATGCTCTTGGAGTAAGACATATCAAAAGAACTGGAGAGCTTCCATTAAGTA 780
Qy 261 ProGlnLeuLysMetGluIle 267
Db 781 CCACAAATTAATAATGGAGATA 801
RESULT 4
US-10-651-722-23
; Sequence 23, Application US/10651722
; Publication No. US20040048302A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE, MP-1
; FILE REFERENCE: D0073 DIV
; CURRENT APPLICATION NUMBER: US/10/651,722
; PRIOR FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 10/067,443
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 1526
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-651-722-23
Alignment Scores:
Pred. No.: 1,32e-169 Length: 1526
Score: 1385.00 Matches: 267
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0
US-10-649-273-2_COPY_148_414 (1-267) x US-10-651-722-23 (1-1526)

```
OY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
DB 1 ATGGAGGCTCATGCACTTACTATTAGGTGACCAATAAAGTAGAATTTCTTTTATTAGTT 60
OY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
DB 61 CTTTGATTTCTGGAGGTCACTGTCTGTGGCAATGATTCAGAGAGTTTCAAGATTTTCTG 120
OY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
DB 121 CTTCTTGGAAAGCTTTGGACATAGCACAGGTGACATGCTTGACAAAGTGGCAAGAGA 180
OY 61 LeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyLysAlaIleGluHis 80
DB 181 CTTTCTTATATAAAACATCCAGAGTCTCCACCATGATGTGTGGAAAGCCATAGAACAT 240
OY 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLys 100
DB 241 TTGGGCCAAACAAGGAATAGATTTCATTTGACATCAAACTCCCTTGCAATGCTAAA 300
OY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
DB 301 AATTGTGATTTTCTTTTACTGACATTCACACGTTACTGATTAATAATTAATGAAG 360
OY 121 GlyLysGluGluGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaAla 140
DB 361 GAAAAAGAGAGGTATTGAGAAAGGGCAAAATCTGTCTTCACAGACAGCATTTGCTGCC 420
OY 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
DB 421 ACAGTACACACACACAAATGCAATGCAATCTGTGAAAGAACACATCGGCTAATCTGTTT 480
OY 161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyVal 180
DB 481 TGTAAACAGAGACCTTGTTACTCTCAAAATTAATGCACTGTTGATCTGGGGTGC 540
OY 181 AlaSerAsnPheTyrlleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThr 200
DB 541 GCAAGTAATCTTCAATCCGACAGCTCTGGAATTTTAAACAAACGAAACACAGTGCAT 600
OY 201 LeuLeuCysProProProArgLeuCysThrAspAsnGlyIleMetIleAlaTrpAsnGly 220
DB 601 TTGTTGTCTCTCTCCACAGATGACATGATTAATGCAATTCATTCATGCAATGCAAT 660
OY 221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleLeuGluIleArgTyrglu 240
DB 661 ATTTGAAAGACTAGTCTGCTGGCATTTTACATGACATAGAAAGGCATCCCTATGAA 720
OY 241 ProLysCysProLeuGlyValAlaAspIleSerLysGluValGlyGluAlaSerIleLysVal 260
DB 721 CCAAAATGTCCTCTTGGAGTGAATATTCATTAAGAAAGTGGAGAGCTTCATTAAGATA 780
OY 261 ProGlnLeuLysMetGluIle 267
DB 781 CCACAATTAATAATGAGATA 801
RESULT 5
US-10-067-443-1
; Sequence 1, Application US/10067443
; Publication No. US20030082782A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE HIGHLY EXPRESSED IN
; FILE REFERENCE: D0073 NP
; CURRENT APPLICATION NUMBER: US/10/067,443
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 1
; LENGTH: 2197
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (231)..(1472)
US-10-067-443-1
Alignment Scores:
Pred. No.: 2,276-169 Length: 2197
Score: 1385.00 Matches: 267
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-10-649-273-2_COPY_148_414 (1-267) x US-10-067-443-1 (1-2197)
OY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
DB 672 ATGGAGGCTCATGCACTTACTATTAGGTGACCAATAAAGTAGAATTTCTTTTATTAGTT 731
OY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
DB 732 CTTTGATTTCTGGAGGTCACTGTCTGTGGCAATGATTCAGAGAGTTTCAAGATTTTCTG 791
OY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
DB 792 CTTCTTGGAAAGCTTTGGACATAGCACAGGTGACATGCTTGACAAAGTGGCAAGAGA 851
OY 61 LeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyLysAlaIleGluHis 80
DB 852 CTTTCTTATATAAAACATCCAGAGTCTCCACCATGATGTGTGGAAAGCCATAGAACAT 911
OY 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLys 100
DB 912 TTGGCCAAACAAGGAATAGATTTCATTTGACATCAAACTCCCTTGCAATGCTAAA 971
OY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
DB 972 AATTGTGATTTTCTTTTACTGACATTCACACGTTACTGATTAATAATGAAG 1031
OY 121 GlyLysGluGluGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaAla 140
DB 1032 GAAAAAGAGAGGTATTGAGAAAGGGCAAAATCTGTCTTGACAGACAGCATTTGCTGCC 1091
OY 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
DB 1092 ACAGTACACACACACAAATGCAATGCAATCTGTGAAAGAACACATCGGCTAATCTGTTT 1151
OY 161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyVal 180
DB 1152 TGTAAACAGAGACCTTGTTACTCTCAAAATTAATGCACTGTTGATCTGGGGTGC 1211
OY 181 AlaSerAsnPheTyrlleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThr 200
DB 1212 GCAAGTAATCTTCAATCCGACAGCTCTGAAATTTTAAACAAACGAAACAGTGCAT 1271
OY 201 LeuLeuCysProProProArgLeuCysThrAspAsnGlyIleMetIleAlaTrpAsnGly 220
DB 1272 TTGTTGTCTCTCTCCACAGATGACATGACATTAATGACATTAATGATTCATGAGATGCT 1331
OY 221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleLeuGluIleArgTyrglu 240
DB 1332 ATTTGAAAGACTAGTCTGCTGGCATTTTACATGACATAGAAAGGCATCCGCTATGAA 1391
OY 241 ProLysCysProLeuGlyValAlaAspIleSerLysGluValGlyGluAlaSerIleLysVal 260
DB 1392 CCAAAATGTCCTCTTGGAGTGAATATTCATTAAGAAAGTGGAGAGCTTCATTAAGATA 1451
OY 261 ProGlnLeuLysMetGluIle 267
```

Db 1452 CCACAAATTAAAAATGAGATA 1472

RESULT 6

US-10-649-273-1

Sequence 1, Application US/10649273

Publication No. US20040043407A1

GENERAL INFORMATION:

APPLICANT: Bristol-Myers Squibb Company

TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE, MP-1

FILE REFERENCE: D0073 CNT

CURRENT APPLICATION NUMBER: US/10/649,273

CURRENT FILING DATE: 2003-08-27

PRIOR APPLICATION NUMBER: US 60/266,518

PRIOR FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: US 10/067,443

PRIOR FILING DATE: 2002-02-05

PRIOR APPLICATION NUMBER: US 60/282,814

PRIOR FILING DATE: 2001-04-10

NUMBER OF SEQ ID NOS: 71

SOFTWARE: PatentIn version 3.2

SEQ ID NO 1

LENGTH: 2197

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (231)..(1472)

US-10-649-273-1

Alignment Scores:

Pred. No.: 2,27e-169 Length: 2197

Score: 1385.00 Matches: 267

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x US-10-649-273-1 (1-2197)

QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20

DB 672 ATGGAGGCTCATGCACTTATCTATTAGTTCACCAATTAAGTAGAATTTCTTTTGGTT 731

QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40

DB 732 CTTTGGATTTCTGGAGGTCACTGCTGTTGGCATTTAGTTCAAGGAGTTTCAGATTTCTG 791

QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 60

DB 792 CTTCTTGGAAAGCTTTTGCATATGACACAGGTGACATGCTTGACAAAGTGGCAAGAGA 851

QY 61 LeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyValAlaIleGluHis 80

DB 852 CTTTCTTAATAAATCATTCAGAGTGTCTCCACATAGTGTGGGAAGCCATAGAAAT 911

QY 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProPheLysHisAlaLys 100

DB 912 TTGGCCAAACAAGAAATAGATTTCATTTTGCATCAACCTCCCTTGCATCATGCTAAA 971

QY 101 AsnGlyAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120

DB 972 AATTGGATTTTCTTTTCTGATGCACTTCAACAGTTACTATTAATAAATAATTAAG 1031

QY 121 GluLysGlnGlnGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 140

DB 1032 GAAAGAGAGAGAGATTTGAGAGGGGCAATCTGCTTCAGAGCAAGACATTTGCTGCC 1091

QY 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160

DB 1092 AAGGTACACACACATGAGATTCATCTTGTGAAAGAACATCGGGGCAATCTGTTT 1151

QY 161 CysLysGlnArgAspLeuProGlnAsnAsnAlaValLeuValAlaSerGlyVal 180

Db 1152 TGTAAAGACAGACACTGTTTACTTCAAAATATATGACAGTACTGTTGCATCTGCTGTCTC 1211

QY 181 AlaSerAspPheThrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr 200

DB 1212 GCAAGTAACTTCTATATCCGAGAGCTCTGGAATTTTAAACAAGCAACAGTGCAT 1271

QY 201 LeuLeuCysProProPheArgLeuCysThrAspAsnGlyIleMetIleAlaIleTrpAsnGly 220

DB 1272 TTGTTGTCTCTCTCCAGACTATGACTGATATATGATCATATGATGATGAGATGCT 1331

QY 221 IleGluArgLeuAlaGlnGlyIleLeuHisAspIleGlnGlyIleArgPheGlu 240

DB 1332 ATTAAAGACTACGTCGCTGCGCATTTTACATGACATAGAAAGGCAATCCGCTATGAA 1391

QY 241 ProLysCysProLeuGlnGlyValAspIleSerLysGlnValGlnGluAlaSerIleLysVal 260

DB 1392 CCATAATGTCCTCTTGGAGTATGACATATCAAAAGATTGAGAGAGCTTCCATTAAGTA 1451

QY 261 ProGlnLeuLysMetGluIle 267

DB 1452 CCACAAATTAAAAATGAGATA 1472

RESULT 7

US-10-651-722-1

Sequence 1, Application US/10651722

Publication No. US20040048302A1

GENERAL INFORMATION:

APPLICANT: Bristol-Myers Squibb Company

TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE, MP-1

FILE REFERENCE: D0073 DIV

CURRENT APPLICATION NUMBER: US/10/651,722

CURRENT FILING DATE: 2003-08-29

PRIOR APPLICATION NUMBER: US 60/266,518

PRIOR FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: US 10/067,443

PRIOR FILING DATE: 2002-02-05

PRIOR APPLICATION NUMBER: US 60/282,814

PRIOR FILING DATE: 2001-04-10

NUMBER OF SEQ ID NOS: 71

SOFTWARE: PatentIn version 3.2

SEQ ID NO 1

LENGTH: 2197

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (231)..(1472)

US-10-651-722-1

Alignment Scores:

Pred. No.: 2,27e-169 Length: 2197

Score: 1385.00 Matches: 267

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x US-10-651-722-1 (1-2197)

QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20

DB 672 ATGGAGGCTCATGCACTTATCTATTAGTTCACCAATTAAGTAGAATTTCTTTTGGTT 731

QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40

DB 732 CTTTGGATTTCTGGAGGTCACTGCTGTTGGCATTTAGTTCAAGGAGTTTCAGATTTCTG 791

QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 60

DB 792 CTTCTTGGAAAGCTTTTGCATATGACACAGGTGACATGCTTGACAAAGTGGCAAGAGA 851

QY 61 LeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyValAlaIleGluHis 80

Db 852 CTTTCTTAATAAATCAATCCAGAGCTGCTCCACCATGAGTGGGAAAGCCATGAACAT 911
Qy 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProLeuHisHisAlaLys 100
Db 912 TTGGCCAAACAGAGAAATGATTTTCAATTTTGACATCAACCTCCCTTGCACTGATGCTAAA 971
Qy 101 AsnGAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
Db 972 AATTGATTTTCTTTTACTGACCTTCAACAGCTTACTATTAATAATTAATGAAG 1031
Qy 121 GlnLysGlnGlnGlyIleGlyLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 140
Db 1032 GAAAAAGAGAGAGATGTAAGAAAGGCGCAATCTGTCTTACACAGACATGTCGCC 1091
Qy 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
Db 1092 ACGATACAGACACCAATGGATCTCATCTTGTGAAGAAACACATCGGGCTATTCTGTTT 1151
Qy 161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyVal 180
Db 1152 TGTAAACAGAGAGACTGTTACTCAAAATTAATGACAGTACTGGTTCATCTGGTGGTGC 1211
Qy 181 AlaSerAsnPheTyrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr 200
Db 1212 GCAAGTAACCTTCTATATCCGACAGAGCTGTGAAATTTTAAACAACGACACAGTGCACT 1271
Qy 201 LeuLeuCysProProProArgLeuCysThrAspAsnGlyIleMetIleAlaTPAsnGly 220
Db 1272 TTGTTTGTCTCTCTCCAGACTAGCATGATTAATGGCATTTATGATGATGAGAAATGGT 1331
Qy 221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGlyGlyIleArgTyrGlu 240
Db 1332 ATTGAAGAAGTACTAGCTGCTGGCATTTTACATTAACATAGAGGCATCCGCTAAGAA 1391
Qy 241 ProLysCysProLeuGlyValAspIleSerLysGlyValGlyGlnAlaSerIleLysVal 260
Db 1392 CCAAAATGTCTCTTGGAGTAGACATATCAAAAGAGTTGAGAGCTTCATTAAGATA 1451
Qy 261 ProGlnLeuLysMetGluIle 267
Db 1452 CCACAAATTAATAATGAGAGATA 1472

RESULT 8
US-10-480-988-36
; Sequence 36, Application US/10480988
; Publication No. US2005069877A1
; GENERAL INFORMATION:
; APPLICANT: GANDHI, Ameena R.; KABLE, Amy E.;
; APPLICANT: SWARNAKAR, Anita; HAPALIA, April J.A.;
; APPLICANT: TRAN, Bao; DUGAN, Brendan M.;
; APPLICANT: WARREN, Bridget A.; ISON, Craig H.;
; APPLICANT: HONCHBL, Cynthia D.; NGUYEN, Daniel B.;
; APPLICANT: LU, Dying, Aina M.; LEE, Ernestine A.;
; APPLICANT: YUE, Henry; FORSYTHE, Ian J.;
; APPLICANT: BARROSO, Ines; RAMKUMAR, Jayalaxmi;
; APPLICANT: GRIFFIN, Jennifer A.; LI, Joana X.;
; APPLICANT: YANG, Junming; THANGAVELU, Kavitha;
; APPLICANT: GIETZEN, Kimberly J.; DING, Li;
; APPLICANT: BAUGHEN, Mariah R.; BOROWSKY, Mark L.;
; APPLICANT: YAO, Monique G.; CHAMLA, Nandinder K.;
; APPLICANT: MASON, Patricia M.; GURURAJAN, Rajagopal;
; APPLICANT: LEE, Sally; BECHA, Shanya D.;
; APPLICANT: LEE, Soo Yeun; TRAN, Uyen K.;
; APPLICANT: ELLIOTT, Vicki S.; LIO, Wen;
; APPLICANT: SPRAGUE, William W.; TANG, Y. Tom;
; APPLICANT: LU, Yan; ZEBARADJIAN, Yeganeh
; TITLE OF INVENTION: PROTEIN MODIFICATION AND MAINTENANCE MOLECULES
; FILE REFERENCE: PF-1040 USN
; CURRENT APPLICATION NUMBER: US/10/480, 988
; PRIOR FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: PCT/US02/19360
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/300,508

; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/303,445
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US 60/305,405
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/311,442
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 60/314,821
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/315,992
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: US 60/378,205
; PRIOR FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 2572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7632424CB1
US-10-480-988-36
Alignment Scores:
Pred. No.: 2,87e-169 Length: 2572
Score: 1385.00 Matches: 267
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-10-649-273-2_COPY_148_414 (1-267) x US-10-480-988-36 (1-2572)
Qy 1 MetGlnAlaHisAlaLeuThrIleArgLeuThrAsnLysValGlnPheProPheLeuVal 20
Db 585 ATGAGAGCTCATGACACTTACTATTAGGTGACCAATTAAGATTTCTTTTATTTGATT 644
Qy 21 LeuLeuIleSerGlyGlnHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
Db 645 CTTTGAATTTCTGGAGGTCACTGTCTGTGGCATTTAGTCAAGAGATTCAGATTTCTG 704
Qy 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
Db 705 CTTCTTGGAAGATCTTTTGACATGACACAGTACATGCTTGACAGAGTGCAAGAGA 764
Qy 61 LeuSerLeuIleLysHisProGlnCysSerThrMetSerGlyGlyLysAlaIleGlnHis 80
Db 765 CTTTCTTAATAAACAATCCAGAGTCTCCACATGAGTGGGAAAGCCATAGAACAT 824
Qy 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLys 100
Db 825 TTGGCCAAACAGAGAAATGATTTTCAATTTTGACATCAAACTCCCTTGACATCATGCTAAA 884
Qy 101 AsnGAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
Db 1085 AATTGATTTTCTTTTACTGACCTTCAACAGCTTACTGATTAATAATTAATGAAG 944
Qy 121 GlnLysGlnGlnGlyIleGlyLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 140
Db 945 GAAAAAGAGAGAGATTTGAGAAAGGCGCAATCTGTCTTACAGACAGACATGTCGCC 1004
Qy 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
Db 1005 ACGATACAGACACCAATGGATGTCATCTGTGAAAAGAACACATCGGGCTATTCTGTTT 1064
Qy 161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyVal 180
Db 1065 TGTAAAGAGAGAGACTGTTACTCAAAATTAATGACAGTACTGGTTCATCTGGTGGTGC 1124
Qy 181 AlaSerAsnPheTyrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr 200
Db 1125 GCAAGTAACCTTCTATATCCGACAGAGCTGTGAAATTTTAAACAACGACACAGTGCACT 1184

Qy 201 LeuLeuCySPProProAArgLeuCySPThraspaSnglyllemetleatpansgly 220
Db 1185 TTGTTGTGTCTCTCCAGAGACTATGATTAAGGATTAATGATGATGAATGAT 1244
Qy 221 TleuLeuAArgLeuAArgLeuGlylleuLeuHisAsp1leuGlylleuGlyArg 240
Db 1245 ATTGAAGAGCTACGTCTGCTGGCTTGGCATTTTACATGACATTAAGGATCCCTATGAA 1304
Qy 241 ProlysCySPProleuGlyValAsp1leuSerlyeGluValGlyGluAsp1leuVal 260
Db 1305 CCAAAATGCTCTTGGAGTACATATCAAAAGAGTTGGAGAGCTTCATTAAGA 1364
Qy 261 ProGlnLeuLeuMetGluile 267
Db 1365 CCACATTTAAATGAGATA 1385
RESULT 9
US-10-649-273-2
; Sequence 21, Application US/10067443
; Publication No. US20030082782A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE HIGHLY EXPRESSED I
; TITLE OF INVENTION: SPINAL CORD, MP-1
; FILE REFERENCE: D0073 NP
; CURRENT APPLICATION NUMBER: US/10/067,443
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 1387
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-649-273-2
US-10-649-273-2
Alignment Scores:
Pred. No.: 3,25e-166 Length: 1387
Score: 1358.50 Matches: 266
Percent Similarity: 91.10% Conservative: 0
Best Local Similarity: 98.09% Mismatches: 1
Query Match: 98.09% Indels: 25
DB: 14 Gaps: 1
US-10-649-273-2_COPY_148_414 (1-267) x US-10-649-273-21 (1-1387)
Qy 1 MetGluAlaHisAlaLeuThr1leArgLeuThrAsnLysValGluPheProPheLeuVal 20
Db 465 ATGAGAGGCTCATGCACTTACTATTAAGTTGACCAATTAAGTAGAATTTCTTTTGTAGTT 524
Qy 21 LeuLeuHisSerGlyGlyHisCySPLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
Db 525 CTTTTCATTTCTGAGAGTCACTCTGTTTGGCATTTAGTTCAAGAGTTTCAAGATTTCTG 584
Qy 41 LeuLeuGlySerLeuAsp1leuAlaProGlyAspMetLeuAspLysValAlaArg 60
Db 585 CTTCTGGAAGCTTTGACATAGCACAGGTGACATGCTTGACAGAGTGAGCAAGAGA 644
Qy 61 LeuSerLeu1leuHisPProGlyCySPSerThreMetSerGlyGlyValAlaGluHis 80
Db 645 CTTTCTTAATAAATCATGAGGTCTCCACCATGAGTGTGGGAAGCCATAGACAT 704
Qy 81 LeuAlaLysGlnGlyAsnArgPheHisPheAsp1leuLysProPheLeuHisHisAlaLys 100
Db 705 TTGGCCAAACAAGGAATAGATTATTTGATTTGACATCAACCTCTTGACATCAAGTAAA 764
Qy 101 AsnCySPAspPheSerTherGlyLeuGlnHisValThrAspLys1leuMetLysLys 120
Db 765 AATGTGATTTTCTTTTACGTGACATTCACACGTTACTGATAAATTAATGAAG 824

Qy 121 GluLysGluGlylle-----GluLysGlyGln1leuSerSerAla 126
Db 825 GAAAGAGAGAGATATATTTCTAATTAGTAAGTTGAACAGATTAATATTCCTGATTG 884
Qy 127 -----GluLysGlyGln1leuSerSerAla 135
Db 885 TGCCATAAATAGCTGCTCATTTCTGAGTATGAGAAAGGGGCAATCTGCTTACACA 944
Qy 136 AlaAsp1leuAlaThrValGlnHisThreMetAlaCySPHisLeuValLysArgThrHis 155
Db 945 GCACACATTTGCTGCCACAGTACAGACACAAATGGCATTCATCTTGTGAAAAGACAT 1004
Qy 156 ArgAla1leuPheCySPGlnArgAspLeuPProGlnAsnAsnAlaValLeuVal 175
Db 1005 CGGCTATTTCTTTTGAACAGAGACTTGTACTCAAAATTAATGAGTACTGCTT 1064
Qy 176 AlaSerGlyGlyValAlaSerAsnPheThr1leArgAlaGluAlaLeuThrAsn 195
Db 1065 GCATCTGTGTGTGCGAAGTAACTTCTATATCCGACAGAGCTCTGGAATTTTAAACAA 1124
Qy 196 AlaThrGlnCySPThreLeuCySPProProAArgLeuCySPThraspaSnglyllemet 215
Db 1125 GCAACACAGTGCACCTTGTGTGTCTCTCCAGACTATGACATTAATGACATTAAG 1184
Qy 216 TleuLeuAArgLeuAArgLeuAArgLeuAArgLeuAArgLeuAArgLeuAArgLeu 235
Db 1185 ATTGATGAGAAATGATTAATGAAGACTAGTGTGCTTGGCATTTTACATGACATGA 1244
Qy 236 Gly1leuArgGlyGluProLysCySPProleuGlyValAsp1leuSerLysGluValGly 255
Db 1245 GGCATCGCTATGAAACCAAAATGCTCTTGGAGTACATATCAAAAGAGTTGAGAA 1304
Qy 256 AlaSer1leuLysValPProGlnLeuLysMetGluile 267
Db 1305 GCTTCAATAAAGTACCAATTAATAATGAGATA 1340
RESULT 10
US-10-649-273-21
; Sequence 21, Application US/10649273
; Publication No. US20040043407A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE, MP-1
; TITLE OF INVENTION: SPINAL CORD, MP-1
; FILE REFERENCE: D0073 CNT
; CURRENT APPLICATION NUMBER: US/10/649,273
; CURRENT FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US 10/067,443
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 1387
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-649-273-21
US-10-649-273-21
Alignment Scores:
Pred. No.: 3,25e-166 Length: 1387
Score: 1358.50 Matches: 266
Percent Similarity: 91.10% Conservative: 0
Best Local Similarity: 98.09% Mismatches: 1
Query Match: 98.09% Indels: 25
DB: 18 Gaps: 1
US-10-649-273-2_COPY_148_414 (1-267) x US-10-649-273-21 (1-1387)
Qy 1 MetGluAlaHisAlaLeuThr1leArgLeuThrAsnLysValGluPheProPheLeuVal 20

Db	465	ATGGAGGCTCACTGACTTACTTATGAGTTGACCAATAAAGTAGAATTTCTTTTAGTT	524
Qy	21	LeuLeuLlIeSerGlyGlyIhIeCySLeuLeuAlaIeuValGlnGlyValSerAspPheLeu	40
Db	525	CTTTTGATTTTCGAGAGTCACTGTCGTGGCATTTAGTTCCAGAGATTTCCAGATTTTCTG	584
Qy	41	LeuLeuGlyIysSerLeuAspLlIeAlaProGlyAspMetLeuAspLysValAlaArgArg	60
Db	585	CTTCTGGAAAGCTCTTTTGACATAGCACAGGTCAGTCCTTACCAAGGTGGCAAGAGA	644
Qy	61	LeuSerLeuLlIeLysHisProGlyCysSerThrMetSerGlyLysValAlaIleGluHis	80
Db	645	CTTTCTTTAATATAAACATCCACAGAGTCTCCACATCAGTGTGTGGAAAGCCATAGAGCAT	704
Qy	81	LeuAlaLysGlnGlyAsnArgPheHisPheAspLlIeLysProProLeuHisSHIaLys	100
Db	705	TTGGCCAAACAAGAGAAATGATTTTCATTTTGACATCAAACTCCCTGTGCATGCTTAA	764
Qy	101	AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysLlIeIleMetLys	120
Db	765	AATGTGATTTTCTTTTACTGACTCAACACGTTACTGATTAATAATTAAGAAGAAAG	824
Qy	121	GluLysGlnGlnGlyLlIe-----GluLysGlyGlnLlIeLeuSerSerAla	135
Db	825	GAAAAAGAGAGAGATATATTTCTAATTAAGTAAAGTGAACAGATAAATATCTCGATTG	884
Qy	127	-----GluLysGlyGlnLlIeLeuSerSerAla	135
Db	885	TGCTCAAAATATGCTGCTCATTTCTTCAGAGTATGAAGAGGGCAAAATCCTGCTTCAGCA	944
Qy	136	AlaAspLlIeAlaIleThrValGlnHisThrMetAlaCysHisIleuValLysArgThrHis	155
Db	945	GCAGACATTTGCTGCCACAGTACAGCACACAAATGCAATGCAATCTTGTGAAAAAACACAT	1000
Qy	156	ArgAlaIleLeuPheCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLleuVal	175
Db	1005	CGCGGTATTTCTTTTGTGTAAGCAGAGACACTTGTACCTCAAAATATAGCAGTACGTGT	1060
Qy	176	AlaSerGlyGlyValAlaSerAsnPheYrIleArgArgAlaLeuGlnLlIeLeuThrAsn	195
Db	1065	GCATGTGTGTGGTGCAGAAATTAATCTTATATCCGACAGACTCTGGAAATTTTAAACAAAC	1120
Qy	196	AlaThrGlnCysThrLeuLeuCysProProProArgLeuCysThrAspAsnGlyLlIeMet	215
Db	1125	GCACACACGTGACTTTGTGTGTCTCTCCACAGCTATGACCTGATATATGCAATTATG	1180
Qy	216	IleAlaTPasngYlIleGluArgLeuArgAlaGlyLeuGlyLlIeLeuHisAspLlIeGlu	235
Db	1185	ATTGCATGGAATGATGATTAAGAAAGATCAAGTGGCTTGGGCAATTTTACATGACATAGAA	1240
Qy	236	GlyIleArgYrGluProLysCysProLeuGlyValAspLlIeSerLysGlyValGlyGlu	255
Db	1245	GGCATCCGCTATGAACCAAAATGTCTCTTTGGAGTAGACATATCAAAAGAAAGTTGGAGAA	1300
Qy	256	AlaSerLlIeLysValProGlnLeuLysMetGluIle	267
Db	1305	GCTTCCATATAAAGTACCACAATTAATAAATGAGAGATA	1340
RESULT 11			
US-10-651-722-21			
Sequence 21, Application US/10651722			
Publication No. US20040048302A1			
GENERAL INFORMATION:			
APPLICANT: Bristol-Myers Squibb Company			
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE, MP-1			
FILE REFERENCE: D0073 DIV			
CURRENT APPLICATION NUMBER: US/10/651,722			
CURRENT FILING DATE: 2003-08-29			
PRIOR APPLICATION NUMBER: US 60/266,518			
PRIOR FILING DATE: 2001-02-05			
PRIOR APPLICATION NUMBER: US 10/067,443			
PRIOR FILING DATE: 2002-02-05			
PRIOR APPLICATION NUMBER: US 60/282,814			

PRIOR FILING DATE: 2001-04-10
 NUMBER OF SEQ ID NOS: 71
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 21
 LENGTH: 1387
 TYPE: DNA
 ORGANISM: homo sapiens
 US-10-651-722-21

Alignment Scores:
 Pred. No.: 3,256-166 Length: 1387
 Score: 1358.50 Matches: 266
 Percent Similarity: 91.10% Conservative: 0
 Best Local Similarity: 91.10% Mismatches: 1
 Query Match: 98.09% Indels: 25
 DB: 18 Gaps: 1

US-10-649-273-2_COPY_148_414 (1-267) x US-10-651-722-21 (1-1387)

QY	1	MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal	20
DB	465	ATGGAGGCGTCACGACCTTAATTGAGTGGACCAATAAGTGAATTTCTTTTAAAGTT	524
QY	21	LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlyGlyValSerAspPheLeu	40
DB	525	CTTTGATTTCTGGAGGTCACTGTCTGTTGGCATTTGTTCAAGAGTTTCAGATTTTCG	584
QY	41	LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg	60
DB	585	CTTCTGGAAAGTCTTTGGACATAGACACAGGTGACATGCTTGCACAAGGTGGCAAGAGA	644
QY	61	LeuSerLeuIleLysHisProGlyCysSerThrMetSerGlyGlyLysAlaIleGluHis	80
DB	645	CTTTCTTAATATAAACATCCAGAGTGTCCACCATAGTGTGGGAAGCCATAGAGCAT	704
QY	81	LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProPheLeuHisAlaLys	100
DB	705	TTGGCCAAACAAGAAATATGATTTCAATTTGACATCAACCTCCCTTGACATACGCTAAA	764
QY	101	AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys	120
DB	765	AATTGTGATTTTCTTTTACGTGACTTCAACACGTTTACATGATTAATAATATGAAGAAAG	824
QY	121	GluLysGluGluGlyIle-----	126
DB	825	GAATAAGAGGAGGATATATTTCTAATATGATAAGTTGAACAGATAATAATTCCTGAGATTG	884
QY	127	-----GluLysGlyGlnIleLeuSerSerAla	135
DB	885	TGCCTAAATAATAGCTGCTCATTTCTGACAGTATGAGAAAGGGGCAAAATCCGTCTTCGCA	944
QY	136	AlaAspIleAlaAlaThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHis	155
DB	945	GCGAGCATTTGCTGCCACAGTACAGACACAAATAGCGATGTCATCTTTGTGAATAAACAACAT	1000
QY	156	ArgAlaIleLeuPheCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuVal	175
DB	1005	CGGGCTAATTCCTGTTTGTATAGACAGAGACTTGTATCTCAAAATATATGACATGCTGTT	1060
QY	176	AlaSerGlyGlyValAlaSerAsnPheTrpIleArgArgAlaLeuGluIleLeuThrAsn	195
DB	1065	GCATCTGGGTGGTGGCAAGTAACTTATATTCGAGAGGCTTGGAAATTTTAAACAAC	1120
QY	196	AlaTrpGlnCysThrLeuLeuCysProProArgLysCysThrAspAsnGlyIleMet	215
DB	1125	GCAACACAGTGCATCTTTGTGTCTCTCTCCCAAGACTATGCACTGATATATGCGATTATG	1180
QY	216	IleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGlu	235
DB	1185	ATTGCATGAAGATGATTAAGAAAGCTACGTGGGCTTGGGCACTTTTACATGACATGGAA	1240
QY	236	GlyIleArgTrpGluProLysCysProLeuGlyValAspIleSerIleGlyValGlyGlu	255


```
Db 1245 GGATCCGCTATGAAACCAAAATGCTCTTGGAGTAGACATATACAAAGAAATTGGAGAA 1304
Qy 256 ALaSerIleYValProGlnLeuLYsMetGluIle 267
Db 1305 GCTTCCATTAAGTACCAACATTAATAAATGGAGATA 1340

RESULT 12
US-10-887-553A-1047
; Sequence 1047, Application US/10887553A
; Publication No. US20050085436A1
; GENERAL INFORMATION:
; APPLICANT: Garza, Dan
; APPLICANT: Li, Hao
; TITLE OF INVENTION: Method to treat conditions associated
; TITLE OF INVENTION: with insulin signalling dysregulation
; FILE REFERENCE: 4-33262
; CURRENT APPLICATION NUMBER: US/10/887,553A
; PRIOR FILING DATE: 2004-07-08
; PRIOR APPLICATION NUMBER: 60/485,883
; NUMBER OF SEQ ID NOS: 1208
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1047
; LENGTH: 1387
; TYPE: DNA
; ORGANISM: human
US-10-887-553A-1047

Alignment Scores:
Pred. No.: 3,25e-166 Length: 1387
Score: 1358.50 Matches: 266
Percent Similarity: 91.10% Conservative: 0
Best Local Similarity: 91.10% Mismatches: 1
Query Match: 98.09% Indels: 25
DB: Gaps: 1

US-10-649-273-2_COPY_148_414 (1-267) x US-10-887-553A-1047 (1-1387)
Qy 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLYsValGluPheProPheLeuVal 20
Db 465 ATGAGAGCTCATGACCTTACTATTAGGTGACCAATAAGATGAAATTCCTTTTCTAGTT 524
Qy 21 LeuLeuIleSerGlyValHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
Db 525 CTTTGTATTTCTGAGAGGCTGCTGCTGTTGGCATTAGTTCAAGAGTTTCAGATTTCTG 584
Qy 41 LeuLeuGlyLYsSerLeuAspIleAlaProGlyAspMetLeuAspLYsValAlaArgArg 60
Db 585 CTTCTTGGAAGGCTCTTGGACATAGCACAGTGACATGCTTGACAAAGTGGCAAGAGA 644
Qy 61 LeuSerLeuIleLYsHisProGlyCysSerThrMetSerGlyLYsAlaIleGluHis 80
Db 645 CTTTCTTTAATRAAACAATCCAGAGTGTCCACACATGAGTGGGAAAGCCATAGAGCAT 704
Qy 81 LeuAlaLYsGlnGlyAsnArgPheHisAspAspIleLYsProProLeuHisHisAlaLYs 100
Db 705 TTGGCCCAACAAAGAAATAGATTTCATTTTGAACATCAACCTCCCTTCATCATCTAA 764
Qy 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLYsIleIleMetLYsLYs 120
Db 765 AATTGTGATTTCTTTTACTGACCTTCAACGCTTACTGATTAATAATATGAAAAAG 824
Qy 121 GlnLYsGlnGlnGlyIle-----GlnLYsGlyGlnIleLeuSerSerAla 135
Db 825 GAAAAAGAGGAGATATTTCTAATTAGTAAAGTTGAACAGATAATATTCCTGATTC 884
Qy 127 -----GlnLYsGlyGlnIleLeuSerSerAla 135
Db 885 TGCTTAATAATAGCTGCTCATTTCTGACAGTATAGAGGCGCAATCTGCTTCACAGA 944
Qy 136 AlaAspIleAlaIleThrValGlnHisThrMetAlaCysHisLeuValLYsArgArgHis 155
Db 945 GCAGACATTCGTCACAGTACAGACACAAATGGCATGATCATCTTGTAAGAAAGACACT 1004
```

```
Qy 156 ArgAlaIleLeuPheCysLYsGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuVal 175
Db 1005 CGGGCTATTCTGTTTGTAAAGCAGAGACCTTGTTACTCCAAAATAATGCACTGCTGTT 1064
Qy 176 ALaSerGlyValAlaSerAsnPheTYrIleArgArgAlaLeuGlnIleLeuThrAsn 195
Db 1065 GCATCTGTGTGTGTCGCAAGTACTTCTATATCCGACAGCTCTGAAATTTTAACAAAC 1124
Qy 196 AlaThrGlnCysThrLeuLeuCysProProProAlaGlyLeuCysThrAspAsnGlyIleMet 215
Db 1125 GCACACAGTGACCTTGTGTGTCTCTCTCCACACTATGACATGATATGCACTTATG 1184
Qy 216 ILeAlaTPAsnGlyIleGlnArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGln 235
Db 1185 ATTGCATGAATGTGATTGAAGAAGCTACAGTGGTGGCTTGCGGCATTTTACATGACATAGA 1244
Qy 236 GlyIleArgTYrGluProLYsCysProLeuGlyValAspIleSerIleGlyValGlyGln 255
Db 1245 GGATCCCGCTATGAAACCAAAATGCTCTTGGAGTAGACATATCAAAAGAAAGTTGGAGAA 1304
Qy 256 ALaSerIleLYsValProGlnLeuLYsMetGluIle 267
Db 1305 GCTTCCATTAAGTACCAACATTAATAAATGGAGATA 1340

RESULT 13
US-10-012-140-6
; Sequence 6, Application US/10012140
; Publication No. US20030009017A1
; GENERAL INFORMATION:
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Kapeller-Livermann, Rosana
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
; TITLE OF INVENTION: 14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
; FILE REFERENCE: 381552004900
; CURRENT APPLICATION NUMBER: US/10/012,140
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/246,768
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,772
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,185
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-140-6

Alignment Scores:
Pred. No.: 6.47e-165 Length: 1245
Score: 1348.00 Matches: 260
Percent Similarity: 98.50% Conservative: 3
Best Local Similarity: 97.38% Mismatches: 4
Query Match: 97.33% Indels: 0
DB: Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x US-10-012-140-6 (1-1245)
Qy 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLYsValGluPheProPheLeuVal 20
Db 442 ATGAGAGCTCATGACCTTACTATTAGTTGACCAATTAAGATTAATTCCTTTTCTAGTT 501
Qy 21 LeuLeuIleSerGlyValHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
Db 502 CTTTGATTTCTGAGAGTCACTGCTGTTGGCATTAGTTCAAGAGATTTTCAGATTTTCTG 561
Qy 41 LeuLeuGlyLYsSerLeuAspIleAlaProGlyAspMetLeuAspLYsValAlaArgArg 60
```

```
Db 562 CTTCTTGGAAGCTTTGACATAGACCAAGTGACATGCTTGACAAAGTGGCAAGAGA 621
Qy 61 LeuSerLeuIlelyshisProgluCyseSerThrmeserlyglylysaalieglnhis 80
Db 622 CTTCTTTATATAAATCATCCAGAGTGCTCCACCATGATGCTGGGAAGCATAGAACAT 681
Qy 81 LeuAlaIyGlnGlysnArghPheHisPheaspIlelyserProleuHisHisAlaIys 100
Db 682 TTGGCCCAACAGAAATATGATTCATTTGACATCAAACTCCCTTGACATCATGCTTAA 741
Qy 101 AasnCyAspPheSerPheThrglyleuGlnHisValThrAspIlylelleMetIylys 120
Db 742 AATTGTGATTTTCTTTTACTGACCTTCAACAGCTTCTATTAATAAATTAAGAAACAGG 801
Qy 121 GlnIyGlnGlyIleGlyIyGlnIleleuSerSerAlaIaaspIleAla 140
Db 802 AAACAAGAGAGAGATTGAGAAAGGGCAATCTGTCTTGAGAGACATTTGCTGCC 861
Qy 141 ThrValGlnHisThrmecAlaCyshisLeuValIySarGlnHisAlaIleuPhe 160
Db 862 ACAGTACAGACACAAATGGCATGTCATCTTGTAAGAAACACATCGGCTATTCGTTT 921
Qy 161 CysIyGlnArghAspIleuProGlnAsnAsnAlaValIleuValAlaSerGlyVal 180
Db 922 TGTAAACAGAGAGACTGTGTAACCTCAAAATATGACATGCTGTGATCTGTGTGTC 981
Qy 181 AlaSerAspPheIyIleArghArghAlaIeugIuIleuThrAsnAlaThrglnCysThr 200
Db 982 GCAAGTAACTTCTATATCCGACAGCTCTGGAATTTTAAACAAGCACAGTGCATCT 1041
Qy 201 LeuLeuCyseProProArghleuCyseThraspAsnGlyIleMetIleAlaIleuPhe 220
Db 1042 TTGTTGTGCTCTCTCCACAGACTATGACATGATATGATGATTAATGATGATGAAATGCT 1101
Qy 221 IleGlnArghleuArghAlaGlyIleuGlyIleleuHisAspIleGlyIyIleArghIyGln 240
Db 1102 ATTGAAGAATAGCTGCTGTGGCTGTGGCATTTTACATGACATAGAGGCAATCCGTAAGAA 1161
Qy 241 ProIyCyseProleuGlyValAspIleSerIyGlnValIyGlnAlaSerIleIyVal 260
Db 1162 CCAAATGTCTCTTGAGATGACATATCAAAAGAGTGGAGAGCTTCATCAATAAGTA 1221
Qy 261 ProGlnLeuIyMetGluIle 267
Db 1222 CCACATTAATAAATGAGATA 1242

RESULT 14
US-10-012-140-4
; Sequence 4, Application US/10012140
; Publication No. US2003009017A1
; GENERAL INFORMATION:
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
; TITLE OF INVENTION: 14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
; TITLE OF INVENTION: THERMOF
; FILE REFERENCE: 381552004900
; CURRENT APPLICATION NUMBER: US/10/012,140
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/246,768
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,772
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,185
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1820
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (146) ... (1390)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) ... (1820)
; OTHER INFORMATION: n = A, T, C or G
US-10-012-140-4

Alignment Scores:
Pred. No.: 1,146-164 Length: 1820
Score: 1348.00 Matches: 260
Percent Similarity: 98.50% Conserved: 3
Best Local Similarity: 97.38% Mismatches: 4
Query Match: 97.33% Indels: 0
Gaps: 0

US-10-649-273-2_copy_148_414 (1-267) x US-10-012-140-4 (1-1820)

Qy 1 MetGlnAlaHisAlaLeuThrIleArghleuThrAsnIyValIleuPheProPheLeuVal 20
Db 587 ATGAGAGCTCAATGACATTAATTAATGAGTGGACCAATAAAGTAATTCCTTTTAACTT 646
Qy 21 LeuLeuIleSerGlyIyHisCysLeuLeuAlaIeugIyValIleuSerAspPheLeu 40
Db 647 CTTTGTGATTTCTGAGAGTCACTGTCTGTGCAATTAAGTTCAGATTTCGATTTCTG 706
Qy 41 LeuLeuGlyIySerLeuAspIleAlaProGlyAspMetLeuAspIyValAlaArghArg 60
Db 707 CTTTGTGAAAGCTCTTGACATATGACACAGTGCATGATGATGATGATGATGATGATGAT 766
Qy 61 LeuSerLeuIlelyshisProgluCyseSerThrmeserlyglylysaalieglnhis 80
Db 767 CTTTCTTTAATAAATCAATCCAGAGTGTCTCCACATGATGATGATGATGATGATGATGAT 826
Qy 81 LeuAlaIyGlnGlysnArghPheHisPheaspIlelyserProleuHisHisAlaIys 100
Db 827 TTGGCCCAACAGAAATATGATTCATTTTACATCAAACTCCCTTGACATCATGCTTAA 886
Qy 101 AasnCyAspPheSerPheThrglyleuGlnHisValThrAspIlylelleMetIylys 120
Db 887 AATTGTGATTTTCTTTTACTGACCTTCAACAGCTTCTATTAATAAATTAATGAAACAGG 946
Qy 121 GlnIyGlnGlyIleGlyIyGlnIleleuSerSerAlaIaaspIleAlaIleuPhe 140
Db 947 AAACAAGAGAGAGATTAATGAGAGGGCAATCTGTCTTGACAGACAGATTCGCTCC 1006
Qy 141 ThrValGlnHisThrmecAlaCyshisLeuValIySarGlnHisAlaIleuPhe 160
Db 1007 ACAGTACAGACACAAATGGCATGTCATCTTGTAAGAAACACATCGGCTATTCGTTT 1066
Qy 161 CysIyGlnArghAspIleuProGlnAsnAsnAlaValIleuValAlaSerGlyVal 180
Db 1067 TGTAAAGAGAGAGACTGTGTAACCTCAAAATATGACATGCTGTGATCTGTGTGTC 1126
Qy 181 AlaSerAspPheIyIleArghArghAlaIeugIuIleuThrAsnAlaThrglnCysThr 200
Db 1127 GCAAGTAACTTCTATATCCGACAGCTCTGGAATTTTAAACAAGCACAGTGCATCT 1186
Qy 201 LeuLeuCyseProProArghleuCyseThraspAsnGlyIleMetIleAlaIleuPhe 220
Db 1187 TTGTTGTGCTCTCTCCACAGCTATGACATGATATGATGATTAATGATGATGATGATGAT 1246
Qy 221 IleGlnArghleuArghAlaGlyIleuGlyIleleuHisAspIleGlyIyIleArghIyGln 240
Db 1247 ATTGAAGAATAGCTGCTGTGGCTGTGGCATTTTACATGACATAGAGGATCGCTATGAA 1306
Qy 241 ProIyCyseProleuGlyValAspIleSerIyGlnValIyGlnAlaSerIleIyVal 260
Db 1307 CCAAATGTCTCTTGAGATGACATATCAAAAGAGTGGAGAGCTTCATCAATAAGTA 1366
Qy 261 ProGlnLeuIyMetGluIle 267
Db 1367 CCACATTAATAAATGAGATA 1387
```

```
RESULT 15
US-10-094-749-400
; Sequence 400, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOMYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 400
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-400

Alignment Scores:
Pred. No.: 9,06e-146 Length: 2208
Score: 1204.00 Matches: 239
Percent Similarity: 89.51% Conservative: 0
Best Local Similarity: 89.51% Mismatches: 4
Query Match: 86.93% Indels: 24
DB: 17 Gaps: 1

US-10-649-273-2_COPY_148_414 (1-267) x US-10-094-749-400 (1-2208)

QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
DB 785 ATGAGAGCTCATGCACTTACTATTAGGTGACCAATTAAGTAGAATTCTTTTAACTT 844
QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
DB 845 CTTTGAATTTCTGAGAGTCACTGCTGTGTGCACTTAGTTCAGAGGTTTCAGATTTTCTG 904
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
DB 905 CTTCTGCAAGAGTCTTTGACATGACACAGGTGACATGCTTGACAGGTGCGAAGAGA 964
QY 61 LeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyLysAlaIleGluHis 80
DB 965 CTTCCTTAAATAAATCCAGAGTGTCTCCACCATGAGTGTGGGAAGCCATAGAACAT 1024
QY 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLys 100
DB 1025 TTGGCCAAACAGGAATAATAGATTTCATTTGACATCAAACTCCCTTCATCATCTAA 1084
QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
DB 1085 AATTGATATTCTCTTTACTGACCTTCACACGTTACGATTAATAATATGAAAAAG 1144
```

```
QY 121 GluLysGluGluGlyIleGluLysGlyClnIleLeuSerSerAlaAlaAspIleAlaAla 140
DB 1145 GAAAAAGAGGAAGGATTTGAGAGAGGCAAAATCTGCTTCAGCAGCAGACATTCCTGCC 1204
QY 141 ThrValGlnHisGlnMetAlaCysHisIleuValLysArgThrHisArgAlaIleLeuPhe 160
DB 1205 ACAGTACGACACACATGGCATGTCATCTTGAAAAAGAACACATCGGGCTAATCTGTTT 1264
QY 161 CysLysGlnArgAspLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyVal 180
DB 1265 TGTAAAGCAGAGAGCTTGTACCTCAAAATAATGACATGCTTGCAATCTGTGTGTGC 1324
QY 181 AlaSerAspPheTyrlleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr 200
DB 1325 GCAAGTAACTTCTGTATCCGACAGAGCTCTGAAAATTTTAAACAAACGCAACAGTGCACT 1384
QY 201 LeuLeuCysProProProArgLeuLysThrAspAsnGlyIleMetIleAlaTrpAsnGly 220
DB 1385 TTGTTGTCTCTCCCTCCAGACTATGCACTGATTAATGCAATTATATGCA----- 1435
QY 221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrlu 240
DB 1435 ----- 1435
QY 241 ProLysCysProLeuGlyValAspIleSerLysGluValGlyGluAlaSerIleLysVal 260
DB 1436 ---TGATCTCTCTTGAGTGAAGTACATATCAAAAGAGTTGGAGAACTTCCATTAAGTA 1492
QY 261 ProGlnLeuLysMetGluIle 267
DB 1493 CCACATTTAAAAATGAGATA 1513
```

Search completed: June 17, 2005, 08:03:19
Job time : 1824.86 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 16, 2005, 15:19:45 ; Search time 2858.39 Seconds
(without alignments)
4051.513 Million cell updates/sec

Title: US-10-649-273-2_COPY_176_414

Perfect score: 1240
Sequence: 1 L1ALVQGVSDFLILGKSLDI.....DISKEVGEASIKVPOLNMEI 239

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame.p2n.model -DEV=x1p
-Q=/cgn2_1/USPTO.spool_p/US10649273/runat_15062005_11416_6030/app_query.fasta_1.1429
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-INITs=bits -START=1 -END=-1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10649273.@CEN_1.1.7742.@runat_15062005_11416_6030 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1240	100.0	1416	6	AR541929 Sequence
2	1240	100.0	1526	6	AR428809 Sequence
3	1240	100.0	1908	9	BC011904 Homo sapi
4	1240	100.0	2197	6	AR428803 Sequence

5	1213.5	97.9	1387	6	AR428808	AR428808 Sequence
6	1213.5	97.9	1387	9	HSN295148	AJ295148 Homo sapi
7	1203	97.0	1245	6	AX664697	AX664697 Sequence
8	1203	97.0	1820	6	AX664695	AX664695 Sequence
9	1086	87.6	1844	10	BC058172	BC058172 Mus muscu
10	1079	87.0	1017	10	BC038910	BC038910 Mus muscu
11	1059	85.4	2208	6	AX713716	AX713716 Sequence
12	1059	85.4	2208	9	AK055441	AK055441 Homo sapi
13	983	79.3	1546	10	BC078974	BC078974 Rattus no
14	944.5	76.2	1109	10	BC051211	BC051211 Mus muscu
15	809	65.2	860	5	BX930694	BX930694 Gallus ga
16	806	65.0	1522	5	BX934991	BX934991 Gallus ga
17	803	64.8	1558	5	BX930863	BX930863 Gallus ga
18	725	58.5	14364	6	AR428807	AR428807 Sequence
19	725	58.5	84115	9	AC013468	AC013468 Homo sapi
20	666	53.7	249601	2	AC114153	AC114153 Rattus no
21	666	53.7	308652	2	AC121478	AC121478 Rattus no
22	646.5	52.1	256751	10	AC129295	AC129295 Mus muscu
23	610	49.2	875	6	CQ721898	CQ721898 Sequence
24	384.5	31.0	1474	3	AK113378	AK113378 Clona int
25	369.5	29.8	121251	5	AL591593	AL591593 Zebrafish
26	356	28.7	20021	10	AF367967S1	AF367967 Mus muscu
27	356	28.7	179252	10	AF131205	AF131205 Mus muscu
28	345	27.8	117322	5	AL672217	AL672217 Zebrafish
29	332.5	26.8	1576	3	AY051882	AY051882 Drosophi
30	332.5	26.8	1601	6	CQ606432	CQ606432 Sequence
31	332.5	26.8	1656	6	CQ606431	CQ606431 Sequence
32	332.5	26.8	14679	2	AC018262	AC018262 Drosophi
33	332.5	26.8	180263	3	AE010671	AE010671 Drosophi
34	332.5	26.8	207432	3	AE003513	AE003513 Drosophi
35	324	26.1	1443	8	AY024338	AY024338 Arabidops
36	324	26.1	1474	8	AY117283	AY117283 Arabidops
37	324	26.1	1567	8	AY063864	AY063864 Arabidops
38	322	26.0	1557	8	AY084577	AY084577 Arabidops
39	311.5	25.1	1672	8	AK070912	AK070912 Oryza sat
40	305	24.6	110000	2	AP065011	AP065011 Sequence
41	262.5	21.2	333800	1	SM591792	SM591792 Sinorhizo
42	260	21.0	1718	8	AK099665	AK099665 Oryza sat
43	257	20.7	349354	1	BX640416	BX640416 Bordetell
44	254	20.5	260271	1	AB017258	AB017258 Molbachi
45	253	20.4	349116	1	AP003003	AP003003 Mesorhizo

ALIGNMENTS

RESULT 1
AR541929 LOCUS AR541929 1416 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 177 from patent US 6743619.
ACCESSION AR541929
VERSION AR541929.1 GI:53934009
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1416)
AUTHORS Tang,Y.T., Zhou,P., Goodrich,R., Liu,C., Asundi,V., Ren,P.,
Zhang,J., Zhao,Q.A., Yang,Y., Xue,A.D., Wehman,T., Wang,J.-R.,
Wang,D. and Drmanac,R.T.
TITLE Nucleic acids and polypeptides
JOURNAL Patent: US 6743619-A 177 01-JUN-2004;
FEATURES
source location/Qualifiers
1..1416
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 2,242-113
Score: 1240.00 length: 1416
Percent Similarity: 100.00% Matches: 239
Best Local Similarity: 100.00% Conservative: 0
Query Match: 100.00% Mismatches: 0
Indels: 0

DB: 6 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x AR541929 (1-1416)

QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlySerLeuAspIle 20
 |||||
 DB 586 CTGTGGCACTTGAAGAGAGTTTCAGATTTCCTTGTGGAAGCTTTGGACATA 645

QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
 |||||
 DB 646 GCACCAAGTGACATGCTTGAACAGGTGGCAAGAGACTTTCTTAATAAATCCAGAG 705

QY 41 CysSerThrMetSerGlyGlyValAlaIleGlnHisLeuAlaLysGlnGlyAsnArgPhe 60
 |||||
 DB 706 TGCTCCACCATGATGCTGGGAAAGCATAGAACATTGGCCAAACAGAAATAGATT 765

QY 61 HisPheAspIleLysProProLeuHisHisAlaLysAsnCyAspPheSerPheThrGly 80
 |||||
 DB 766 CATTTTGACATCAAACTCCCTTGCAATGCTAAATAATGATTTTCTTTTACTGGA 825

QY 81 LeuGlnHisValThrAspLysIleIleMetLysLysGlnLysGlnGlyIleGluLys 100
 |||||
 DB 826 CTTCACACGTTACTGATTAATAATATATGAAAAAGAAAAAGGAGGATTTGAGAG 885

QY 101 GlyGlnIleLeuSerSerAlaAlaAspIleAlaIleThrValGlnHisThrMetAlaCys 120
 |||||
 DB 886 GGGCAATCTGTCTTCAAGACAGCATTTGCTGCCACAGTACAGCACAAATGGCATGT 945

QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
 |||||
 DB 946 CATCTTGTAAGAAAGAACATCGGCTATTCTGTTTGTATGACAGAGACTTGTACT 1005

QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyValAlaSerAsnPheThrIleArgArg 160
 |||||
 DB 1006 CAAATTAATGCGATCTGTGTCATCTGTGTGTGCGAAGTACTTATATCCGAGA 1065

QY 161 AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
 |||||
 DB 1066 GCTCTGGAATTTTAACAAACGACACAGTGCATTTGTGTGTCTCTCCACAGCTA 1125

QY 181 CysThrAspAsnGlyIleMetIleAlaTPAsnGlyIleGluArgLeuArgAlaGlyLeu 200
 |||||
 DB 1126 TGACCTGATTAATGCGATTAATGATTCATGGAATGATTAAGACTACGTCGGCTTG 1185

QY 201 GlyIleLeuHisAspIleGlnGlyIleArgTyrgLupProLysCysProLeuGlyValAsp 220
 |||||
 DB 1186 GGCATTTTAACATGACATAGAGGCAATCCGCTATGAACCAAAATGTCTCTTGGAGTAGAC 1245

QY 221 IleSerLysGluValGlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 239
 |||||
 DB 1246 ATATCAAAAGAGTTGAGAGAGCTTCATTAAGTACCAATTAATAATGAGATA 1302

RESULT 2
 AR428809 1526 bp DNA linear PAT 18-DEC-2003
 LOCUS Sequence 23 from patent US 6642041.
 ACCESSION AR428809
 VERSION AR428809.1 GI:40188595
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1526)
 AUTHORS Chen,J., Feder,J.N., Nelson,T.C., Krystek,S.R. and Duclos,F.
 TITLE Polynucleotides encoding a novel metalloprotease, MP-1
 JOURNAL Patent: US 6642041-A 23 04-NOV-2003;
 FEATURES location/Qualifiers
 source 1..1526
 /organism="Unknown"
 /mol_type="genomic DNA"

Pred. No.: 2,456-113 Length: 1526
 Score: 1240.00 Matches: 239
 Percent Similarity: 100.00% Conservative: 0
 Best local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x AR428809 (1-1526)

QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlySerLeuAspIle 20
 |||||
 DB 85 CTGTGGCACTTGAAGAGAGTTTCAGATTTCCTTGTGGAAGCTTTGGACATA 144

QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
 |||||
 DB 145 GCACCAAGTGACATGCTTGAACAGGTGGCAAGAGACTTTCTTAATAAATCCAGAG 204

QY 41 CysSerThrMetSerGlyGlyValAlaIleGlnHisLeuAlaLysGlnGlyAsnArgPhe 60
 |||||
 DB 205 TGCTCCACCATGATGCTGGGAAAGCATAGAACATTGGCCAAACAGAAATAGATT 264

QY 61 HisPheAspIleLysProProLeuHisHisAlaLysAsnCyAspPheSerPheThrGly 80
 |||||
 DB 265 CATTTTGACATCAAACTCCCTTGCAATGCTAAATAATGATTTTCTTTTACTGGA 324

QY 81 LeuGlnHisValThrAspLysIleIleMetLysLysGlnLysGlnGlyIleGluLys 100
 |||||
 DB 325 CTTCACACGTTACTGATTAATAATATATGAAAAAGAAAAAGGAGGATTTGAGAG 384

QY 101 GlyGlnIleLeuSerSerAlaAlaAspIleAlaIleThrValGlnHisThrMetAlaCys 120
 |||||
 DB 385 GGGCAATCTGTCTTCAAGACAGCATTTGCTGCCACAGTACAGCACAAATGGCATGT 444

QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
 |||||
 DB 445 CATCTTGTAAGAAAGAACATCGGCTATTCTGTTTGTATGACAGAGACTTGTACT 504

QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyValAlaSerAsnPheThrIleArgArg 160
 |||||
 DB 505 CAAATTAATGCGATCTGTGTCATCTGTGTGTGCGAAGTACTTATATCCGAGA 564

QY 161 AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
 |||||
 DB 565 GCTCTGGAATTTTAACAAACGACACAGTGCATTTGTGTGTCTCTCCACAGCTA 624

QY 181 CysThrAspAsnGlyIleMetIleAlaTPAsnGlyIleGluArgLeuArgAlaGlyLeu 200
 |||||
 DB 625 TGACCTGATTAATGCGATTAATGATTCATGGAATGATTAAGACTACGTCGGCTTG 684

QY 201 GlyIleLeuHisAspIleGlnGlyIleArgTyrgLupProLysCysProLeuGlyValAsp 220
 |||||
 DB 685 GGCATTTTAACATGACATAGAGGCAATCCGCTATGAACCAAAATGTCTCTTGGAGTAGAC 744

QY 221 IleSerLysGluValGlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 239
 |||||
 DB 745 ATATCAAAAGAGTTGAGAGAGCTTCATTAAGTACCAATTAATAATGAGATA 801

RESULT 3
 BC011904 1908 bp mRNA linear PRI 23-DEC-2003
 LOCUS Homo sapiens O-sialoglycoprotein endopeptidase-like 1, mRNA (cDNA
 DEFINITION clone MGC:20293 IMAGE:4121450), complete cds.
 ACCESSION BC011904
 VERSION BC011904.2 GI:40225818
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1908)
 Straubeberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Schenker,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buettner,K.H., Schaefer,C.F., Bhat,N.K.,

Alignment Scores:

TITLE	JOURNAL	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov On Dec 19, 2003 this sequence version replaced gi:15080281. Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nigrl.nih.gov Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R., Machuro,Q.L., Masiello,C., Maserati,B., Mastrian,S.D., McCloskey,J.C., McCowell,J., Pearson,R., Stantirpop,S., Thomas,P.J., Touchman,J.W., Tsurgoun,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.	NIH-MGC Project URL: http://mgc.nci.nih.gov On Dec 19, 2003 this sequence version replaced gi:15080281. Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nigrl.nih.gov Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R., Machuro,Q.L., Masiello,C., Maserati,B., Mastrian,S.D., McCloskey,J.C., McCowell,J., Pearson,R., Stantirpop,S., Thomas,P.J., Touchman,J.W., Tsurgoun,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.	NIH-MGC Project URL: http://mgc.nci.nih.gov On Dec 19, 2003 this sequence version replaced gi:15080281. Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nigrl.nih.gov Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R., Machuro,Q.L., Masiello,C., Maserati,B., Mastrian,S.D., McCloskey,J.C., McCowell,J., Pearson,R., Stantirpop,S., Thomas,P.J., Touchman,J.W., Tsurgoun,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.	NIH-MGC Project URL: http://mgc.nci.nih.gov On Dec 19, 2003 this sequence version replaced gi:15080281. Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nigrl.nih.gov Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R., Machuro,Q.L., Masiello,C., Maserati,B., Mastrian,S.D., McCloskey,J.C., McCowell,J., Pearson,R., Stantirpop,S., Thomas,P.J., Touchman,J.W., Tsurgoun,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.	NIH-MGC Project URL: http://mgc.nci.nih.gov On Dec 19, 2003 this sequence version replaced gi:15080281. Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nigrl.nih.gov Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R., Machuro,Q.L., Masiello,C., Maserati,B., Mastrian,S.D., McCloskey,J.C., McCowell,J., Pearson,R., Stantirpop,S., Thomas,P.J., Touchman,J.W., Tsurgoun,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.	NIH-MGC Project URL: http://mgc.nci.nih.gov On Dec 19, 2003 this sequence version replaced gi:15080281. Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nigrl.nih.gov Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R., Machuro,Q.L., Masiello,C., Maserati,B., Mastrian,S.D., McCloskey,J.C., McCowell,J., Pearson,R., Stantirpop,S., Thomas,P.J., Touchman,J.W., Tsurgoun,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.	NIH-MGC Project URL: http://mgc.nci.nih.gov On Dec 19, 2003 this sequence version replaced gi:15080281. Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nigrl.nih.gov Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R., Machuro,Q.L., Masiello,C., Maserati,B., Mastrian,S.D., McCloskey,J.C., McCowell,J., Pearson,R., Stantirpop,S., Thomas,P.J., Touchman,J.W., Tsurgoun,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.	NIH-MGC Project URL: http://mgc.nci.nih.gov On Dec 19, 2003 this sequence version replaced gi:15080281. Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nigrl.nih.gov Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R., Machuro,Q.L., Masiello,C., Maserati,B., Mastrian,S.D., McCloskey,J.C., McCowell,J., Pearson,R., Stantirpop,S., Thomas,P.J., Touchman,J.W., Tsurgoun,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.	

misc_feature		215..1111		/gene="ROSPEPL1"		/note="Peptidase_M22; Region: Glycoprotease family"		/db_xref="CDId:pfam00814"																									
ORIGIN																																	
<p>Alignment Scores:</p> <table border="1"> <thead> <tr> <th>Pred. No.:</th> <th>3..19e-113</th> <th>Length:</th> <th>1908</th> </tr> </thead> <tbody> <tr> <td>Score:</td> <td>1240.00</td> <td>Matches:</td> <td>239</td> </tr> <tr> <td>Percent Similarity:</td> <td>100.00%</td> <td>Conservative:</td> <td>0</td> </tr> <tr> <td>Best local Similarity:</td> <td>100.00%</td> <td>Mismatches:</td> <td>0</td> </tr> <tr> <td>Query Match:</td> <td>100.00%</td> <td>Indels:</td> <td>0</td> </tr> <tr> <td>DB:</td> <td>9</td> <td>Gaps:</td> <td>0</td> </tr> </tbody> </table>										Pred. No.:	3..19e-113	Length:	1908	Score:	1240.00	Matches:	239	Percent Similarity:	100.00%	Conservative:	0	Best local Similarity:	100.00%	Mismatches:	0	Query Match:	100.00%	Indels:	0	DB:	9	Gaps:	0
Pred. No.:	3..19e-113	Length:	1908																														
Score:	1240.00	Matches:	239																														
Percent Similarity:	100.00%	Conservative:	0																														
Best local Similarity:	100.00%	Mismatches:	0																														
Query Match:	100.00%	Indels:	0																														
DB:	9	Gaps:	0																														
US-10-649-273-2_COPY_176_414 (1-239) x BC011304 (1-1908)																																	
QY	1	LeuLeuAlaLeuValGInGIyValSerAspPheLeuLeuLeuGIyLySerLeuAspIle	20																														
DB	629	CTGTTGGCATTAAGTTCAAGAGGTTTCAGATTTTCGCTTCTTGAAAGTCTTTGGACATA	688																														
QY	21	AlaProGIyAspMetLeuAspLyValAlaArgArgLeuSerLeuIleLyHisProGIu	40																														
DB	689	GCACGAGGTGACATGCTTGACCAAGGTGCCAAGAAAGACATTTCTTAATTAACATCCAGAG	748																														
QY	41	CysSerThrMetSerGIyGIyLyValAlaIleGIuHisLeuAlaLySGInGIyAsnArgPhe	60																														
DB	749	TGCTTCCACCAAGAGGTGGGAAAGCCATAAACCATTGGCCCAACAGAAATGATTT	808																														
QY	61	HisPheAspIleLyProProLeuHisHisAlaLyAsnCyAspPheSerPheThrGIy	80																														
DB	809	CATTTTGACATCAAACTCCCTTCATCATCTTAAATTTGTGATTTTCTTTTACTGGA	868																														
QY	81	LeuGIuHisValThrAspLyIleIleMetLyLySGInLyGIyIleGIuLyS	100																														
DB	869	CTTCAACACGTTACGATATAATATATATGAAAAAGAAAGGAAAGGTATGAGAGAG	928																														
QY	101	GIyGIuIleLeuSerSerAlaAlaAspIleAlaIleThrValGIuHisThrMetAlaCyS	120																														
DB	929	GGCCAAATCCGTGCTTCAGCAGCAGACATTCCTCCACAGATACGACACACATGGCATGT	988																														
QY	121	HisLeuValLySarGIuHisArgAlaIleLeuPheCyLySGInArgAspLeuLeuPro	140																														
DB	969	CATCTTGGAAGAAAGACATCGGGCTATTCTGTTTGTATAGCAGAGACATGTATCTT	1048																														
QY	141	GIuHisAsnAlaValLeuValAlaSerGIyGIyValAlaSerAsnPheLyIleArgArg	160																														
DB	1049	CAAAATATATGACGTACTGTTGCATCTGTGTGTGTCGCAAGATTACTTATATATCGCAGA	1108																														
QY	161	AlaLeuGIuIleLeuThrAsnAlaThrGIuCySerThrLeuLeuCyProProProArgLeu	180																														
DB	1109	GCTCTGGAATATTTTAAACAAACGCAACAGTGCACCTTGTGTGTGTCCTCTCCAGACTA	1166																														
QY	181	CysThrAspAsnGIyIleMetIleAlaIleArgAsnGIyIleGIuArgLeuArgAlaGIyLeu	200																														
DB	1169	TGCACGTATTAATGCGATTATATGATTCAGATGGAATGTATTGAAAGCTACGTCTGCTTG	1222																														
QY	201	GIyIleLeuHisAspIleGIuGIyIleArgTyGIuProLyCyAspProLeuGIyValAsp	220																														
DB	1229	GGCATTTTATACATGACATAGAGGCGCATCGCTATGCAACAAATATGCTCTTGAGATGAC	1288																														
QY	221	IleSerLySGInValGIyGIuAlaSerIleLyValProGIuLeuLyMetGIuIle	239																														
DB	1289	ATATCAAAAGAGATTGGAGAGCTTCCATTAAGATACCAATTAATAAATGGAGATA	1345																														

RESULT 4

AR428803

LOCUS

DEFINITION

Sequence 1 from patent US 6642041.

2197 bp

DNA

linear

PAT 18-DEC-2003

ACCESSION AR428803
VERSION AR428803.1 GI:40188589
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2197)
AUTHORS Chen, J., Feder, J.N., Nelson, T.C., Krystek, S.R. and Duclos, F.
TITLE Polynucleotides encoding a novel metalloprotease, MP-1
JOURNAL Patent: US 6642041-A 1 04-NOV-2003;
FEATURES
source
1. 2197
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 3,77e-113 Length: 2197
Score: 1240.00 Matches: 239
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6
US-10-649-273-2_COPY_176_414 (1-239) x AR428803 (1-2197)
QY 1 LeuLeuAlaLeuValGInGlyValSerAspPheLeuLeuGlySerLeuAspIle 20
DB 756 CTTGTTGGCATTAGTTCAGAGAGTTTCAGATTCTGCTTGGAAAGCTTTGGACATA 815
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 816 GCACCAAGGAGACAGTGTGCAAGGTGGCAAGAACTTTCTTAATAAATCCAGAG 875
QY 41 CysSerThrMetSerGlyGlyValAlaIleGluHisLeuAlaLysGInGlyAsnArgPhe 60
DB 876 TCTCTCCACCATGAGTGTGGAAAGCCATAGAACATTGGCCAAACAGAAATAGATT 935
QY 61 HisPheAspIleLysProPheLeuHisAlaLysAsnCyAspPheSerPheThrGly 80
DB 936 CATTTCAGACATCAACCTCCCTTCAGCATCATGCTAAATATGATTTTCTTTACTGGA 995
QY 81 LeuGlnHisValThrAspLysIleIleMetLysLysGluLysGluGlyIleGluLys 100
DB 996 CTTTCAACAGCTTACTGATTAATAATATGAAAAAGAAAAAGAGAGATTGAGAG 1055
QY 101 GInGInIleLeuSerSerAlaAlaAspIleAlaAlaThrValGInHisThrMetAlaCys 120
DB 1056 GGGCAATCTCTGCTTCAGCAGACATTTGCTGCCACAGTACAGCACAAATGGCATGT 1115
QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
DB 1116 CATCTTGTGAAAAAGAACATCGGGCTATCTGTTTGTAAAGCAGAGACTGTTACT 1175
QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyValAlaAsnSerPheTyrlleArgArg 160
DB 1176 CAATAATATGCACTACTGTTGCACTGTGTGTGTCAGAGTAACTTCTATATCCGAGA 1235
QY 161 AlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
DB 1236 GCTCTGGAATTTTAAACAAGCAGACACAGTGCCTTTGTGTCTCTCCAGACTA 1295
QY 181 CysThrAspAsnGlyIleMetIleAlaTPAsnGlyIleGluArgLeuAlaGlyLeu 200
DB 1296 TGCACATGATATGCACTTATGATTCAGTGAATGATTTGAAAAAGACTACGCTGGCTTG 1355
QY 201 GInIleLeuHisAspIleGluGlyIleArgTyrlleProLysCysProLeuGlyValAsp 220
DB 1356 GGCATTTCATATACATAGAGAGCATCCGCTATGAAACCAAAATGCTCTTGGAGTAC 1415
QY 221 IleSerLysGluValGInGlyAlaSerIleLysValProGlnLeuLysMetGluIle 239
DB 1416 ATATCAAAAGAACTTGGAGAGCTTCATATAAGTACACATTAAAAATGAGAGATA 1472

RESULT 5
AR428808
LOCUS AR428808 1387 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 21 from patent US 6642041.
ACCESSION AR428808
VERSION AR428808.1 GI:40188594
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1387)
AUTHORS Chen, J., Feder, J.N., Nelson, T.C., Krystek, S.R. and Duclos, F.
TITLE Polynucleotides encoding a novel metalloprotease, MP-1
JOURNAL Patent: US 6642041-A 21 04-NOV-2003;
FEATURES
source
1. 1387
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 9.37e-111 Length: 1387
Score: 1213.50 Matches: 238
Percent Similarity: 90.15% Conservative: 0
Best Local Similarity: 90.15% Mismatches: 1
Query Match: 97.86% Indels: 25
Gaps: 1
DB: 6
US-10-649-273-2_COPY_176_414 (1-239) x AR428808 (1-1387)
QY 1 LeuLeuAlaLeuValGInGlyValSerAspPheLeuLeuGlySerLeuAspIle 20
DB 549 CTTGTTGGCATTAGTTCAGAGAGTTTCAGATTCTGCTTGGAAAGCTTTGGACATA 608
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 609 GCACCAAGTGCATGCTTGCACAGGTGGCAAGAACTTTCTTAATAAATCCAGAG 668
QY 41 CysSerThrMetSerGlyGlyValAlaIleGluHisLeuAlaLysGInGlyAsnArgPhe 60
DB 669 TGTCTCCACCATGAGTGTGGAAAGCCATAGAACATTGGCCAAACAGAAATAGATT 728
QY 61 HisPheAspIleLysProPheLeuHisAlaLysAsnCyAspPheSerPheThrGly 80
DB 729 CATTTCAGACATCAACCTCCCTTCAGCATCATGCTTAATATGATTTTCTTTACTGGA 788
QY 81 LeuGlnHisValThrAspLysIleIleMetLysLysGluLysGluGlyIle----- 98
DB 789 CTTCAACAGCTTACTGATTAATAATATGAAAAAGAAAAAGAGAGATATTTGTA 848
QY 98 ----- 98
DB 849 ATTAGTAAAGTTGAACAGATTAATATTCCTGATGTCCTTAATAATAGCTCATTTTC 908
QY 99 -----GluLysGluIleLeuSerSerAlaAlaAspIleAlaAlaThrValGln 115
DB 909 TGCAGTATGAGAGGGGCAATCTGCTTCAGCAGACAGCATTTGCGCACAGTACAG 968
QY 116 HisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGln 135
DB 969 CACCAATGCGATTCATCTTGTGAAAAAGAACATCGGGCTATCTGTTTGTATAGCAG 1028
QY 136 ArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyValAlaAsn 155
DB 1029 AGAGACTTGTACCTCAAAATAATATGCACTGATCGGTGCATCGGTGGTCAAGTAAC 1088
QY 156 PheTyrlleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCys 175
DB 1089 TTTCTATATCCGACAGAGCTCGGAAATTTTAAACAAGCAGACACAGTGCATTTGTTGT 1148
QY 176 ProProProArgLeuCysThrAspAsnGlyIleMetIleAlaTPAsnGlyIleGluArg 195

Dn		1149	CCTCCTCCACAGCATGACTGCATGTAAAGTGGATTATGATTGCATGAATGTAATGAAGA	1208
Oy		196	LeuarAgaIyleuglyl1leuhHisapillegluglyl1eaqfyrGlunProLyCyS	215
Dn		1209	CTAGCGTAGTGCGCTGGCGCATTTTCATGACATGACATGAGCAATCCGCTATGAACAATAATGT	1268
Oy		216	ProleuglyVaIarpiIeserlyagVlaJcylguiaaserijelyeValProGIdeu	235
Dn		1269	CCTTGTGAGTAGACATATCAAAGAGTGGAGAGCTTCATMAAGTACCACATTA	1328
Oy		236	LysMetcGuile	239
Dn		1329	AAAATGAGAGTA	1340
RESULT 6				
LOCUS	HSA295148	1387 bp	mRNA	linear PRI 30-OCT-2000
DEFINITION	Homo sapiens mRNA for putative sialoglycoprotease type 2.			
ACCESSION	AJ295148			
VERSION	AJ295148.1 GI:11071726			
KEYWORDS	metallopeptidase; sialoglycoprotease.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE				
AUTHORS	Chen,J.M., Fortunato,M. and Barrett,A.J.			
TITLE	Cloning and sequencing of a second human putative sialoglycoprotease homologue			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 1387) Chen,J.M.			
AUTHORS	Direct Submission Submitted (27-OCT-2000) Chen J.M., MRC Molecular Enzymology Laboratory, The Babraham Institute, Babraham, Cambridge, CB2 4AT, UNITED KINGDOM			
JOURNAL				
FEATURES				
SOURCE	Location/Qualifiers 1..1387 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /tissue_type="uterus" 24..1343 /function="metallopeptidase of family M22" /codon_start=1 /product="putative sialoglycoprotease type 2" /protein_id="CAC14666.1" /db_xref="GI:11071727" /_db_xref="GOB:O9N4B0" /_db_xref="UniProt/TREMBL:Q9N4B0" /translation="MLILTKTAGVFPEKSKRYVEFLRNFPHBTYFLAKIVIGIEIT SCDPAADVDETVGLSEAIHSOTEEVLKTGVLPAPAOOLHEENIORIVOELASLAS SGSDLSAIRATTIKPGLSLSIGVLSLGVSISLVNQKKPFIPIHMEAHALTIPTNN VERPVVILIIGHCILAVOCVSDFLLKESIDLDARDMDLYARRLSLIKPECST MSGVALEHLAKQGNRRFPDIKPPLHTANKCDFSTFGIOHTTDKIIMKEGESBIPLIL SKVGQINDIPGLCLIKIAHFCKYEKGQILLSSADLAATVOHTMACHLVETRIAILFCPC ORDLIPONNNAVIVAGSVANSFYIRRALEILTNAQCTLCPPEPLCTDNGIMTAMNG IERLRGGLIHLDIEGINRYEPKCPIGVDISKEVGEASIKVPQLKMEI"			
CDS				
ORIGIN				
Alignment Scores:				
Pred. No.:	9,37e-111	Length:	1387	
Score:	1213.50	Matches:	238	
Percent Similarity:	90.15%	Conservative:	0	
Best Local Similarity:	90.15%	Mismatches:	1	
Query Match:	97.86%	Indels:	25	
Bd:	9	Gaps:	1	
US-10-649-273-2_COPY_176_414 (1-239) x HSA295148 (1-1387)				
Oy		1	leueuaialauvalghnglyValSerappPreleuauLeuglyllysSerIeuaspille	20
Dn		549	CTGTGGCAATTAAGTTCAGAGATTTCAGATTTCGTCTTGGAGAAAGCTTTGGACATTA	608

Oy	21	AlaProGlyAspMetLeuAspIysValAlaArgArgIleuSerLeuIleIysHisProGlu	40
Db	609	GCACCGAGTGCATCGTTGACCAAGGTGCAGAAAGACTTCTTTAATAATAAACATCCAGAG	668
Oy	41	CysSerThrMetSerGlyGlyValAlaIleGluHisLeuAlaIysGlnIysAsnArgPhe	60
Db	669	TGCTTCCACATGATGTGTGGGAAGCCATAGACATTTGGCCAAACAAAGAAATAGATT	728
Oy	61	HisPheAspIleIysProProLeuHisHisAlaIysAsnCysAspPheSerPheThrGly	80
Db	729	CATTGTGACATCAAACTCCCTTSCATCATGCTAAATAATTGTGATTTTCTTTATCGGA	788
Oy	81	LeuGlnHisValThrAspIysIleIleMetIysIleGlyGluGluGlyIle-----	98
Db	789	CTTCAACACGTTACTGATTAATAATATAGAAAAAGAAAAAGAGAAAGATATTTCTA	848
Oy	98	-----	98
Db	849	ATTAGTAAAGTTGAACAGATAAATATTCCTGGATTGCTGCTAAATAATAGCTTCATTC	908
Oy	99	-----GluysGlyGlnIleLeuSerSerAlaIleAspIleAlaIleAlaThrValGln	115
Db	909	TGCAGCATATGGAAGGGCCAAATCTGTCTTCACAGACAGACATTCCTGCCACATGAC	968
Oy	116	HisThrMetAlaCysHisLeuValIysArgThrHisArgAlaIleLeuPheCysIysGln	135
Db	969	CACACATGGCATGTCACTCTTGAAAAAGAACACATCGGGCTAATCTGTTGTTATACAG	1028
Oy	136	ArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsn	155
Db	1029	AGAGCTTGTTAACTCCAAATAAAGACAGTACTGTTGATCTGGGTGGTGCAGAAATAC	1088
Oy	156	PheThrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuIleCys	175
Db	1089	TTCTATATCCCAAGAGCTCTGAAATTTTACAAACGACACACAGTGCATTTGTTGTGT	1148
Oy	176	ProProProAlaGlyLeuCysThrAspAsnGlyIleMetIleAlaATPAsnGlyIleGluArg	195
Db	1149	CCTCTCCCAACATATGCATGATATATGCGATTTGATGATGATGATGAATGGATTAAGA	1208
Oy	196	LeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGluProIysCys	215
Db	1209	CTACGTGTGCTGTGGGCAATTTTACATGACATAGAGGCATCCGCTATGAACCAAAATGT	1268
Oy	216	ProLeuGlyValAspIleSerIysGluValGlyGluAlaSerIleIysValProGluLeu	235
Db	1269	CCTCTTGAGATGACATATCAAAAGAGTTGGAGAGCTTCATATAAAGTACCACAATTA	1328
Oy	236	LysMetGluIle	239
Db	1329	AAAAATGAGATA	1340
RESULT 7			
LOCUS	AX664697	1245 bp	DNA
DEFINITION	Sequence 6 from Patent WO02074960.		linear
ACCESSION	AX664697		
VERSION	AX664697.1		GI:29164457
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE			
AUTHORS	leibyl, K.R., Kapeller-Libermann, R. and Glucksmann, M.		
TITLE	1		
JOURNAL	38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions		
FEATURES	of human proteins and uses thereof.		
source	Patent: WO 02074960-A & 24-Sep-2002.		
	Millennium Pharmaceuticals, Inc. (US)		
	Location/Qualifiers		
	1..1245		
	/Organism="Homo sapiens"		

ORIGIN /mol_type="unassigned DNA"
/db_xref="taxon:9606"

Alignment Scores:

Pred. No.:	9,1e-110	Length:	1245
Score:	1203.00	Matches:	232
Percent Similarity:	98.33%	Conservative:	3
Best Local Similarity:	97.07%	Mismatches:	4
Query Match:	97.02%	Indels:	0
	6	Gaps:	0

US-10-649-273-2_COPY_176_414 (1-239) x AX664697 (1-1245)

```
Oy 1 LeuLeuAlaLeuValGInglyValSerAspPheLeuLeuGlyLeuSerLeuAspIle 20
Db 526 CTGTTGGCATTAGTTCAGAGAGTTTCAGATTTCCTGCTTCTTGAAAGCTTTGGACATA 585
Oy 21 AlaProGlyAspMetLeuAspIlyValAlaArgArgLeuSerLeuIleYshIspProGlu 40
Db 586 GCACCGAGTGACATGCTTGCACAGGTGGCAAGAACTTTCTTAATAAATCCAGAG 645
Oy 41 CysSerThrMetSerGlyGlyValAlaIleGlnHisLeuAlaIlySGInglyValArgPhe 60
Db 646 TGCCTCCACATAGTGTGGTGGAAAGCCATAGAACTTTGGCCAAACAGAAATAGATT 705
Oy 61 HisPheAspIleLeuSerProPheLeuHisIshIshIshIshIshIshIshIshIshIsh 80
Db 706 CATTGTGACATCAAACTCCCTTGCATCATGCTAATAATTTGATTTTCTTTTACGGA 765
Oy 81 LeuGlnHisValThrAspIlyValIleIleMetIlySGIulysGluGluGlyIleGluIlys 100
Db 766 CTTCAACACGTTACTGATTAATAATATGAAACAGCAACAGAAAGGATTTGAGAG 825
Oy 101 GlyGlnIleLeuSerSerAlaIshIshIshIshIshIshIshIshIshIshIshIshIsh 120
Db 826 GGGCAATCTCTGCTTCAGACAGACATTTGCTCCACAGTACAGCACAAATGGCATGT 885
Oy 121 HisLeuValIlyArgThrHisArgAlaIleLeuPheCysIlySGInglyValArgPheLeuPro 140
Db 886 CATCTTGTAAAGAAACATCGGGCTATTCTGTTTGTAAAGCAAGAACTTTGTTACT 945
Oy 141 GlnAsnAsnAlaValLeuValAlaSerGlyIlyValAlaSerAsnPheYrIleArgArg 160
Db 946 CAAATAATGACAGTACGTGTTGATCATGCGTGTGGCAAGTAACTTCTATATCCGAGA 1005
Oy 161 AlaLeuGluIleLeuThrAsnAlaIshIshIshIshIshIshIshIshIshIshIshIsh 180
Db 1006 GCTCTGGAATTTTAAACAAACGACACAGTGCATTTGTGTGTCTCTCCACAGACTA 1065
Oy 181 CysThrAspAsnGlyIleMetIleAlaIshIshIshIshIshIshIshIshIshIshIsh 200
Db 1066 TGCACATGATATGCAATGATTTGATTCAGTAAAGTATTAAGACTACGTCGCTTG 1125
Oy 201 GlyIleLeuHisAspIleGluGlyIleArgIlyGluProIlyCysProLeuGlyValAsp 220
Db 1126 GGCATTTCATACATACATAGAGCAATCCGCTATGAAACAAATCTCTCTGAGTGAC 1185
Oy 221 HisSerIlySGIulysAlaIlyGluAlaSerIleIlyValProGlnLeuIlySmetGluIle 239
Db 1186 ATATCAAAAGAGTTGAGAAAGCTTCATTAAGTACCAATTAATAATGAGAGATA 1242
```

RESULT 8
AX664695 LOCUS AX664695 1820 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 4 from Patent WO2074960.
ACCESSION AX664695
VERSION AX664695.1 GI:29164455
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Leiby,K.R., Kapeller-libermann,R. and Glucksmann,M.
TITLE 38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions
of human proteins and uses thereof
JOURNAL Patent: WO 02074960-A 4 26-SEP-2002;
Milleium Pharmaceuticals, Inc. (US)
FEATURES
source Location/Qualifiers
1..1820
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

CDS

146..1390
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD80044.1"
/translation="MLITRTAGVFPFKSRKRYEPLRSFNFHPTLHKLVIET
SCDDTAAAVDETENVGEALHSQTEVHLKTKGIVPPAQQLEHNTORIYQNALAS
GVSPSDLSAIAITTKPGALSLVGLSPSLQVGLKPELPIHMEHALITRLTK
VEPFLVLLISGHCILALVOGVSDFLLKSLDIAGDMIDKARSLIKPECT
MSGKALIEHLAKQNRFPDIPKPLHAKNCDSPFTGLQHTDKXNRKKEGIEKG
OISSADIAATVOHTMAHLVKRTHRAIIEPCXRDLLPQNNAVLVASGVASNYIR
RALEITLMAQCITLCPRLCTDNGIMIANGTBRLAIGLILHIEIGIYEPKPL
GVDSKSGEASIVPOLKMET"

ORIGIN

Alignment Scores:

Pred. No.:	1,43e-109	Length:	1820
Score:	1203.00	Matches:	232
Percent Similarity:	98.33%	Conservative:	3
Best Local Similarity:	97.07%	Mismatches:	4
Query Match:	97.02%	Indels:	0
	6	Gaps:	0

US-10-649-273-2_COPY_176_414 (1-239) x AX664695 (1-1820)

```
Oy 1 LeuLeuAlaLeuValGInglyValSerAspPheLeuLeuGlyLeuSerLeuAspIle 20
Db 671 CTGTTGGCATTAGTTCAGAGAGTTTCAGATTTCCTGCTTCTTGAAAGCTTTGGACATA 730
Oy 21 AlaProGlyAspMetLeuAspIlyValAlaArgArgLeuSerLeuIleYshIspProGlu 40
Db 731 GCACCGAGTGACATGCTTGCACAGGTGGCAAGAACTTTCTTAATAAATCCAGAG 790
Oy 41 CysSerThrMetSerGlyGlyValAlaIleGlnHisLeuAlaIlySGInglyValArgPhe 60
Db 791 TGCCTCCACATAGTGTGGTGGAAAGCCATAGAACTTTGCTCCACAGTACAGCAATG 850
Oy 61 HisPheAspIleLeuSerProPheLeuHisIshIshIshIshIshIshIshIshIshIsh 80
Db 851 CATTGTGACATCAAACTCCCTTGCATCATGCTAATAATTTGATTTTCTTTTACGGA 910
Oy 81 LeuGlnHisValThrAspIlyValIleIleMetIlySGIulysGluGluGlyIleGluIlys 100
Db 911 CTTCAACACGTTACTGATTAATAATATGAAACAGCAACAGAAAGGATTTGAGAG 970
Oy 101 GlyGlnIleLeuSerSerAlaIshIshIshIshIshIshIshIshIshIshIshIshIsh 120
Db 971 GGGCAATCTCTGCTTCAGACAGACATTTGCTCCACAGTACAGCACAAATGGCATGT 1030
Oy 121 HisLeuValIlyArgThrHisArgAlaIleLeuPheCysIlySGInglyValArgPheLeuPro 140
Db 1031 CATCTTGTAAAGAAACATCGGGCTATTCTGTTTGTAAAGCAAGAACTTTGTTACT 1090
Oy 141 GlnAsnAsnAlaValLeuValAlaSerGlyIlyValAlaSerAsnPheYrIleArgArg 160
Db 1091 CAAATAATGACAGTACGTGTTGATCATGCGTGTGGCAAGTAACTTCTATATCCGAGA 1150
Oy 161 AlaLeuGluIleLeuThrAsnAlaIshIshIshIshIshIshIshIshIshIshIshIsh 180
Db 1151 GCTCTGGAATTTTAAACAAACGACACAGTGCATTTGTGTGTCTCTCCACAGACTA 1210
Oy 181 CysThrAspAsnGlyIleMetIleAlaIshIshIshIshIshIshIshIshIshIshIsh 200
```

	DB	1211	TGCACGTATATGGCATTCATTCATTCATGACGAATGGATGGATTAAGAAACATACGCTGGCTGTG	1270
QY	201	GLYILEUHSAPPILEGUIGLYILEKRGYRGIUPPOLYSCYSPROLEUGLYVALASP	220	
DB	1271	GGCATTTTACATGCATATGAAGGCAATCCGCTATGACCAACCAATATGCTCTTGGAGTAC	1330	
QY	221	IIeSerlysgIuValGlyGluAlaSerIIleIysValProGlnLeuIysMetGluIle	239	
DB	1331	ATATCAAAAGAAAGTTGGAGAGCTTCCTATAAAGTACACCAATTTAAAAATGGAGATTA	1387	
RESULT 9				
BC058172				
LOCUS	BC058172	1844 bp	mRNA	linear
DEFINITION	Mus musculus cdna clone	MGC:67870	IMAGE:5012054	complete cds
ACCESSION	BC058172			
VERSION	BC058172.1	GI:34849663		
KEYWORDS	MGC			
SOURCE	Mus musculus	(house mouse)		
ORGANISM	Mus musculus			

REFERENCE
AUTHORS
Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G.,
1 (bases 1 to 1844)
Eumariota, Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Submitted (15-Sep-2003) National Institutes of Health, Mammalian

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk

Email: cgapbbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing By: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brookes, S.,
Dietsch, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P.,
Haddu, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Mahuro, O.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McGowen, L., Pearson, K., Stancicrop, S., Thomas, P.J., Touchman, J.W.,
Tsurgou, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

FEATURES
SOURCE
ClonE distribution: MGC clonE distribution information can be found through the I.M.A.G.E. Consortium/INTL at: <http://image.llnl.gov>
Series: IRAK Plate: 123 Row: d Column: 22
This clonE was selected for full length sequencing because it passed the following selection criteria: matched mRNA gl: 21312463.
Location/Qualifiers
1..1814 ..MGC: mgs0112463

CDS

misc_feature

ORIGIN

Alignment Scores:	
Pred. No.:	6.02e-98
Score:	1086.00
Percent Similarity:	92.05%
Best Local Similarity:	85.77%
Query Match:	87.58%
DB:	10
	Gaps: 0
	Indels: 0
	Matches: 205
	Conservative: 15
	Mismatches: 19
	Length: 1844

US-10-649-273-2_COPY_176_414 (1-239) X BC058172 (1-1844

Oy	1	LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuLeuGlyLysSerLeuAspIle	20
Dd	747	CTGTTGGCATTAGTCCAGAGTGTTCCGATTTCCTGCTCCCTTGGGAAGCTTTGGACATA	806
Oy	21	AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu	40
Dd	807	GCGCCAGGCGGACAGTCTTGACAAAGGTGGCAAGAAAGACTTTTATATCAACATCCAGAA	866
Oy	41	CysSerThrMetSerGlyGlyLysAlaIleGluHisLeuAlaLysGlnGlyAsnArgPhe	60
Dd	867	TGTTCTCAATAGAGTGGGAAAGAGCTATAGAACATTGGCCCAAGACGGAATAGATTTC	926
Oy	61	HisPheAspIleLysProProLeuHisIleHisAlaLysAsnGlyAspPheSerPheThrGly	80
Dd	927	CATTTTACTATCAATCCACCTATACAGAGACTAGAGATTTCGATTTTCTTTTAAACGGA	986
Oy	81	LeuGlnHisValThrAspLysIleIleMetLysLysGluLysGluGluGlyIleGluLys	100
Dd	987	CTTCAACATATTACTAGATACCTATATACACACAAAGAAAGAAAGAAAGGCAATTGAGAG	1046
Oy	101	GlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaIleThrValGlnHisCysIleCys	120
Dd	1047	GGGCAATCTCTGTCATACGCTCCAGACACTTCTGCTGGCGTATACAGCATCAACACGTCGC	1106
Oy	121	HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuPro	140

Db	Accession	Source	Organism	Reference	Authors
Db	1107	CACCTTGCACAAAAAACAACACANTCGTGCTATTCTGTTTTCGACAGCAGAAAAATTTCGTACT			1166
Qy	141	GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPheTyrIleArgArg			160
Db	1167	CCAGCTAACCCAGCATTTAGTTGTTGTCGACAGTGTTCGACGTAACCTTGATCATCTCGAAAA			1226
Qy	161	AlaIleuGluIleIleuThrAsnAlaThrGlnCysThrIleuLeuCysProProPArgIleu			180
Db	1227	GCATTGGAATATGTCGCAAAATGCAACGACGACGTGCTGTGTCTCCCTTCGCAAGACTG			1286
Qy	181	CysThrAspAsnGlyIleMetIleAlaIleTyrAsnGlyIleGluArgLeuArgAlaGlyLeu			200
Db	1287	TGCATCTGACATGACATCATCATATTGTCGATCGAATGGAATGAAGATTACGTGCTGCTG			1346
Qy	201	GlyIleIleuHisAspIleGluGlyIleArgTyrGluProIleCysProLeuGlyValAsp			220
Db	1347	GGCGCTTTTACATGATGTATGAAGACATCCGATATGACCAAAATATGCTCCCTTGGAATGAC			1406
Qy	221	IleSerTyrGluValIleGlyGluAlaSerIleTyrValProGlnIleuIleCysMetGluIle			239
Db	1407	ATATCCAGAGAGATGTCAGAAAGCTGCCATYAAAAGTACCGCATTTAAAAATGACACTT			1463
RESULT 10					
LOCUS	BC038910	1017 bp	mRNA	linear	ROD 21-OCT-2003
DEFINITION	Mus musculus O-sialoglycoprotein endopeptidase-like 1, mRNA (CDNA				
ACCESSION	BC038910				
VERSION	clone IMAGE:5053559), partial cds.				
KEYWORDS	BC038910.1 GI:24433548				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 1017)				
	Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klauer, S.R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Sculter, G.D., Altschul, N.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marsina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S., Carninci, P., Prange, C., Kaha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, X., Gibbs, R.A., Vialleton, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shenchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Buckenfield, Y.S., Krzywinski, M.I., Skalska, J., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
MEDLINE	22388257				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 1017)				
AUTHORS	Straussberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-OCT-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgabbs-rc@mail.nih.gov Tissue Procurement: Jeffrey R. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequencing Centre,				

BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

Steven Jones, Jennifer Aasno, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letlicia Hsiao, Martin Krzywnicki, Reta Kutsche, Oliver Lee, Soo San Lee, Victor Ling, Carrie Mathewson, Candace McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Parabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Nataaja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/IML at: <http://image.llnl.gov>

Series: IRAX Plate: 86 Row: f Column: 12.

Location/Qualifiers

1. 1017

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone_image="5053559"

/tissue_type="Liver, normal. 5 month old male mouse."

/clone_lib="NCI CGAP L19"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

ORIGIN	
Alignment Scores:	
Pred. No.:	1.48e-97
Score:	1079.00
Percent Similarity:	91.63%
Best Local Similarity:	85.77%
Query Match:	87.02%
DB:	10
Length:	1017
Matches:	205
Conservative:	14
Mismatches:	20
Indels:	0
Gaps:	0

US-10-649-273-2_COPY_176_414 (1-239) X BC038910 (1-1017)

```

QY      1 LeuLeuAlaIeValGIgNGlyValIserAspPheLeuLeuLeuGlyysSerLeuAspIle 20
Db      110 CTGTTGGCATTAAGTCAAGAGTGTTCGATTTCCGATTCCTGGGAAGTCTTTGGACATA 169
QY      21 AlaProGlyAspMetLeuAspLysValAlaIArgArgLeuSerLeuIleYshsProGlu 40
Db      170 GCACCAAGCGACATCTCTGCACAGGTGGCAAGAAGACTTTCTTTATCCAACTCCAGAA 229
QY      41 CysSerTherMetSerGlyGlyLysAlaIleGluHisLeuAlaLysGlnGlyAsnArgPhe 60
Db      230 TGTTCACATAGAGTGGTGGAAAGACTATAGAACAGTTGGCCAAAGACGAATATGATTC 289
QY      61 HisPheAspIleLysProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGly 80
Db      290 CATTTTACTATCAATCCACCTATGACAGATGCTAAGAATTCGATTTTCTTTCACGGGA 349
QY      81 LeuGlnHisValaThrAspLysIleIleMetLysGluLysGluGluGlyIleGluLys 100
Db      350 CTTCAACATATTACGATGATGACTAATACACACACAGAAAAGAAAGGCAATTGAGAG 409
QY      101 GlyGlnIleLeuSerSerAlaIaAspIleAlaIaThrValGlnHisThrMetAlaCys 120
Db      410 GGGCAAAATCTCTCATCACTGACCTGACACATCTGCTGCGCGGTACAGCATGACAAGCTGC 469
QY      121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
Db      470 CACCTTGGCAAAAAGAACACATCGCGCTATTCTGTTTGGACAGCAAAAATTTGCTCTCT 529
QY      141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPhenylrIleArgArg 160
Db      530 CCAGCTAACGCGAGTATGATTGATCTGGAGGTTGGCAAGTAACTGTACATCCGAAA 589
QY      161 AlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
Db      590 GCATTGGAATTTGTGCANATCCAAAGCAGTGACCTGTTGTGTCCACTCCAAAGCTG 649

```

QY	181	Cythrisspangylilemetilaltrpnsnlyileguaygleuargalaglileu	200		
Db	650	TGCATGACAAATGCGATCATCATATTCGATGGAATGGAATGAAGATTACGTCGCTGG	709		
QY	201	Glyileleuhiisaaplilegunglyileargtryglnuprolyscysproleuglyvalasp	220		
Db	710	GGCGTTTAACTAGATGTAAGAACATCCGATATGACCAAAATGTCCTTGAGATGAC	769		
QY	221	IlleserlysglyvalglyglualiserlylelysvaIProginleuylsmetGluile	239		
Db	770	ATATCCAGAGAAATGGCAGAAAGCTGCCATAAAGTAAACCGCATTAATAATGACCTT	826		
RESULT 11					
LOCUS	AX1713716	2208 bp	DNA linear PAT 15-APR-2003		
DEFINITION	Sequence 400 from Patent EP1293569.				
ACCESSION	AX1713716				
VERSION	AX1713716.1	GI:29888642			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Iisogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S., Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R., Tanachika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and Masuho,Y.				
TITLE	Full-length cDNA				
JOURNAL	Patent: EP 1293569-A 400 19-MAR-2003.				
FEATURES	Helix Research Institute (JP) ; Research Association for Biotechnology (JP)				
SOURCE	location/Qualifiers				
	1..2208				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Alignment Scores:					
Pred. No.:	3,57e-95	Length:	2208		
Score:	1059.00	Matches:	211		
Percent Similarity:	88.28%	Conservative:	0		
Best Local Similarity:	88.28%	Mismatches:	4		
Query Match:	85.40%	Indels:	24		
DB:	6	Gaps:	1		
US-10-649-273-2_COPY_176_414 (1-239) x AX1713716 (1-2208)					
QY	1	leuleuhalaleuvalgnglyvalserapheleuleuleuglylyserleuaspille	20		
Db	869	CTGTGGGATTAGTTCAAGAGATTTCACATTTTCGCTTGGAAAGCTTTGGACATA	928		
QY	21	AlaProglyaspmetleuaspysvalalargargleuSerleuilelyshisProglu	40		
Db	929	GCACCAAGTGACATGCTTGCACAGGTGGCAGAGACATTCCTTTAATTAACATCCAG	988		
QY	41	CysSerthmetSerglylylylsalaleguluhiSeuualalySGnglyYasnargphe	60		
Db	989	TGCTCCACATGAGTGGTGGGAAACCATPAGAACATTTGGCCAAACAGAAATGATTT	1046		
QY	61	HispheaspilleyProProleuhiishiSalalybaanCysaspPheSerPheThrgly	80		
Db	1049	CATTTCACATCAAACTCCCTTCATCATATGCTAAATAATTGATTTTCCTTTTCTGGA	1108		
QY	81	LeuglnhisvalThaspysileilemetlyslsglysglnuglyylleglulys	100		
Db	1109	CTTCAACACGTTCATCGATTAATAATTAATGAAAAAGAAAAAGAGAAAGTATTGAGAG	1166		
QY	101	GlyglniileuSerSerAlaAlaaspililealalatrvalglnhiThmerAlaCys	120		
Db	1169	GGGCAAAATCTGTCTTCAGCGAGACATGTGCGACAGTACAGACACAAATGGCATGT	1226		

[illegible]

Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuno, Y., Nagai, K. and Isogai, T., NEDO human cDNA sequencing project

Unpublished
3 (bases 1 to 2208)
Isogai, T., Otsuki, T. and Sugiyama, T.
Direct Submission
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 252-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' - 3' - end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
source location/Qualifiers
1..2208
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FEBRA2004592"
/tissue_type="brain"
/clone_id="FEBRA2"
/dev_stage="Fetus"
/note="cloning vector: pMB18SFL3"
344..1438
/note="unnammed protein product"
/codon_start=-1
/protein_id="BAB70923.1"
/db_xref="GI:16550167"
/translation="MLITKTYGVFFPKSRKRYVFLRSFNFHGLTFLKIVGIETSCDDTAAVVDERTGVGEAHSOTVEHLVTGTVPPAQQOLHRENTORLVOEALMSGVSPDLASAIATTIKPGLALSLGVLSFSLQVGLKKPPIHHEMRALTTIRTKVERPRIVLISGGHCLALVQVSDPLILKSLDIDAGDMLDKVARLPILTKPECTMSGGKAIENHAKQNNPFDPKIPPLHAKKDCPPFGLQVHTKTIIMKKKEBEGTQILSSADINATVQHTMACHLVKRTTHRAIIFCKORDLIPQNNAVLVASGVASNFCTRRALEILTNATQCTILCPPLRLCTDNGIMIA"

ORIGIN
Alignment Scores:
Pred. No.: 3,57e-95 Length: 2208
Score: 1059.00 Matches: 211
Percent Similarity: 88.28% Conservative: 0
Best Local Similarity: 88.28% Mismatches: 4
Query Match: 85.40% Indels: 24
DB: 9 Gaps: 1

US-10-649-273-2_COPY_176_414 (1-239) x AK055441 (1-2208)

QY 1 LeuLeuAlaLeuValGInGInValSerASPheLeuLeuGInGInValSerLeuAspIle 20
DB 869 CTTGTGGCATTAGTTCAGGAGGTTTCAGATTTCTGCTTCGAAAGCTTTGGACACTA 928
QY 21 AlaProGInYAspMetLeuAspIleValAlaArgArgLeuSerLeuIleYSHIeProGIn 40
DB 929 GCACCGAGTGACATGCTTGACCAAGTGCGCAAGAACTTCCTTAATAAACAATCCAGAG 988
QY 41 CysSerThrMetSerGInGInValAlaIleGluHisLeuAlaValGInGInValAsnArgPhe 60
DB 989 TCGTCCACCATGAGTGGTGGAAAGCATAGAACATTGGCCAAACAGAAATAGATT 1048
QY 61 HisPheAspIleYsProProLeuHisHisAlaValAsnGInYAspPheSerPheThrGIn 80
DB 1049 CATTGTGACATCAAACTCCCTTGACATCAAGCTAAAAATGTGATTTTCTTTACTGGA 1108
QY 81 LeuGlnHisValThrAspIleIleIleMetIleYsGInGInGInGInGInGInGInGIn 100

DB 1109 CTTCAACAGCTTACTGATTAATAATGAAAAAGCAAGAGGATTCAGAG 1168
QY 101 GInGInIleLeuSerSerAlaAlaAspIleAlaAlaThrValGInHisThrMetAlaCys 120
DB 1169 GGGCAATTCCTGCTTCAGCAGCAGACATTCCTGCAAGTACAGACATGCAATG 1228
QY 121 HisLeuValIleYsArgThrHisArgAlaIleLeuPheCysIleYsGInArgAspLeuPro 140
DB 1229 CATCTTGAAAAAACAACATCGGGCTATCTGCTTTGTAAGCAGAGACTGTTTACT 1288
QY 141 GlnAsnMetAlaValIleValAlaSerGInGInValAlaSerAspPheTyrIleArgArg 160
DB 1289 CAATAATATGAGTACGCTGGTTCATCTGGTGTGCGAAGTAACTTGTATCCGCA 1348
QY 161 AlaLeuGInIleLeuThrAsnAlaThrGInCysThrLeuLeuCysProProArgPhe 180
DB 1349 GCTTGAAATTTTATACAAAGCAACAGTGCATTTGTGTGTCTCTCCGACACTA 1408
QY 181 CysThrAspAsnGInIleMetIleAlaAlaTrpAsnGInIleGInArgLeuArgAlaGInLeu 200
DB 1409 TGCACGTATATGACATATGATTCGA----- 1435
QY 201 GInIleLeuHisAspIleGInGInIleArgTyrGlnProCysProLeuGInValAsp 220
DB 1436 -----TGATGCTCTTGGAGTAC 1456
QY 221 ILeSerIleGInValGInGInAlaSerIleIleYsValProGInLeuYsMetGInIle 239
DB 1457 ATATCAAAAGAGTGGAGAGCTTCATTAAGTACCAATTAATGAGATTA 1513

RESULT 13
BC078974 1546 bp mRNA linear ROD 03-AUG-2004
LOCUS BC078974
DEFINITION Rattus norvegicus cDNA clone IMAGE:7111906, partial cds.
ACCESSION BC078974
VERSION BC078974.1 GI:50926879
KEYWORDS
SOURCE
ORGANISM Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
AUTHORS
1 (bases 1 to 1546)
Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhac, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scherz, T.E., Brownstein, M.J., Udell, T.B., Toibiyoki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.D., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huliyil, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Scherchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalka, U., Smalls, D.E., Schercher, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1546)
Director MGC Project.
DIRECTOR MGC PROJECT.
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
DIRECT SUBMISSION
Submitted (02-AUG-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Cell, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK
NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT

Contact: MGC help desk
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Howard Jacobs
 CDNA Library Preparation: Express Genomics
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 182 Row: f Column: 6
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, Similarity but not identity to protein.
 Location/Qualifiers

FEATURES

CDS

```

1..1546
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:7111906"
/tissue_type="Testis, rat (Brown Norway)"
/clone_lib="NIH MGC_237"
/lab_host="DH10B"
/note="Vector: pExpress1"
144..>1546
/codon_start=1
/product="unknown (protein for IMAGE:7111906)"
/protein_id="AAH78974.1"
/db_xref="GI:50926880"
/transcript="MMLSTKAGATPPPSNVGFRFRVQRALEPHNGLVGIET
SCDYLAAVDETNVGEALHSQTEVLTGGLVPPAQOQHEENQRIVEALAS
GSPSDLSAATTKPGIALSGVLSFVQVLPQPKPPIHMEHALTITLTK
VGPFVLISGCHLALVQSVDFLGLSLDIPDMDKVARSLLKHESCT
MSGKAEHLAKENRPHFTINPMONAKNDPEPTGLOHTDLITHKKEGK
OIISSADIAAIVHATACHAKRTHAIIIFCOONLSPANAVLVSGVSNLYR
RALEIVANATQCTLCPPRLCTDNGIMIANGLERLAKAGLILHVEDIRPEKAGE
IVSMRLVLTALTELDLSVSTHTVAHSPNSGRGANIOTSMCCTOTVVRVHTL
NINERKSKKKKKKKK"

```

ORIGIN

Alignment Scores:

Pred. No.: 8.27e-88 Length: 1546
 Score: 983.00 Matches: 188
 Percent Similarity: 92.06% Conservative: 9
 Best Local Similarity: 87.85% Mismatches: 17
 Query Match: 79.27% Indels: 0
 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x BC078974 (1-1546)

```

QY 1 LeuLeuAlaLeuValGInGlyValISerAspPheLeuLeuGlyLysSerLeuAsp1le 20
DB 669 CTGTGGCGGTAGTGGAGAGGTTTCAGATTTTCCTCGCGAAGTCCCTGAGCAATA 728
QY 21 AlaProGlyAspMetLeuAspLysValAlaIArgArgLeuSerLeuIleLysHisProGlu 40
DB 729 GCGCCAGGCGAGCTTGACAGAGGTGCGAAGAGACTTTCTTTAATCAACATCCAGAA 788
QY 41 CysSerThrMetSerGlyGlyLysAlaIleGluHisLeuAlaLysGInGlyAsnArgPhe 60
DB 789 TGTTCATCATGAGTGGTGGAGAAAGCTATAGAACTTTGGCCAAAGAGAAATGATTC 848
QY 61 HisPheAspIleLysProProLeuHisHisAlaLysAsnGlyAspPheSerPheThrGly 80
DB 849 CACTTACTACTCATCCACCCATGACAGATGCTAAGAACTGTGATTTTCTTTTACGGGA 908
QY 81 LeuGlnHisValIThrAspLysIleIleMetLysLysGluLysGluGluGly1leGluLys 100
DB 909 CTTCAACATGTCACCGATAGCTAATTAACACACAGAAAGAAAGGACATGAGAG 968

```

```

QY 101 GlyGlnIleLeuSerSerAlaAlaAsp1leAlaAlaThrValGlnHisThrMetLacys 120
DB 969 GGGCAAAATCCTGTATCATCGCCGACAGACATTCCTGCGGTACACAGACCAACACGTCGC 1028
QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
DB 1029 CACCTTGGAAAGAAACACATCGCTCTTCTGTTTCCACGACAAATAATTTGCTATCT 1088
QY 141 GlnAsnAlaValLeuValAlaSerGlyValAlaLysAspPheThrLysArg 160
DB 1089 CCACTTAACGACATTTAGTTGTCTGTCTGAGGTGTTCAGAGTACTTGTACATCCGAGA 1148
QY 161 AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
DB 1149 GCATTGAAATTTGTGCAAAATGCAACACATGCACTTTGTGTGTCCTCCGACAGCTG 1208
QY 181 CysThrAspAsnGlyLysMetLysAlaIleArgLysGluArgLeuArgAlaLysLeu 200
DB 1209 TGCACTGCAAAATGATCATGATGATTCATGATGAAATGAAAGATTACGTGCTGCTTG 1268
QY 201 GlyIleLeuHisAspIleGluGlyLysArgThrGluProLys 214
DB 1269 GCATTTCATGATGATGAGAGACATCCGATTCGAAACCAAG 1310

```

RESULT 14

BC051211

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK
 COMMENT
 Contact: MGC help desk
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Marcello Bento Soares, Ph.D.
 CDNA Library Preparation: M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadan@systemsbio.org
 Anup Madan, Jessica Fahy, Erin Helton, Mark Ketteman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 113 Row: b Column: 1
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES

source

```

1..1109
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1327545"
/tissue_type="Thymus gland, mouse"
/clone_lib="Soares_thymus_2dbmt"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac"

```

ORIGIN

Alignment Scores:
 Pred. No.: 3,726-84 Length: 1109
 Score: 944.50 Matches: 183
 Percent Similarity: 82.52% Conservative: 20
 Best Local Similarity: 74.39% Mismatches: 32
 Query Match: 76.17% Indels: 11
 DB: 10 Gaps: 2

US-10-649-273-2_COPY_176_414 (1-239) x BC051211 (1-1109)

```

QY 2 LeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspPheAla 21
DB 148 TTGAAACTGACTGTCGCAATGAGGCTTATCTTGTCTTAAT-----AATTTGCAG 198
QY 22 ProGlyAspMetLeuAspLys-----ValAlaArgArgLeu 33
DB 199 CCAAAATATATTTATGAAAAAATGTAATGCTTTTAAATAGGTGGCAGAAAGACTT 258
QY 34 SerLeuIleLysHisProGlyCysSerThrMetSerGlyLysValAlaIleGluHisLeu 53
DB 259 TCTTTATCAAAACATCCAGATGTTCTACCAATGATGAGTGGAAAAAGCTATAGAACACTTG 318
QY 54 AlaValGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLysAsn 73
DB 319 GCCAAAGAGCGAAATGATTCATTTTATCATCAATCCACTATGACGAATGCTAAGAAAT 378
QY 74 CysAspPheSerPheThrGlyLeuGlnHisLeuValThrAspLysIleIleMetLysValGlu 93
DB 379 TGGCATTTTCTTTCACGGGACTTCACATATATTCATATGAGCTAATTAACACACAAAGAA 438
QY 94 LysGlnGluGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaThr 113
DB 439 AAAGAAAGAGGACTTGAAGAGGCAAAATCTGTCAATCACTGACAGCATTTGCTGCGG 498
QY 114 ValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPheCys 133
DB 499 GTACAGACATGCAACACGTCGCTGCACTTGCAGAAAAAGAACACATCGCGCTATTCGTTTGC 558
QY 134 LysGlnArgAspLeuLeuProGlnHisAsnAlaValLeuValAlaSerGlyGlyValAla 153
DB 559 AAGCAAAAATTTGCTCTCTCCAGCTTAACGCAAGTATTAAGTTATCTTGAAGGTGGCA 618
QY 154 SerAsnPheThrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeu 173
DB 619 AGTAATTTGTATCATCCGAAAGCATTTGGAATTTGCGCAATTCGCAACGCTGCACTTG 678
QY 174 LeuCysProProArgLeuCysThrAspAsnGlyIleMetIleAlaIleArgGlyIle 193
DB 679 TTGTGTCACCTCCAGACTGCTGACATGCAATGCGCATCATGATGCAATGGAATT 738
QY 194 GluArgLeuArgAlaLeuGlyLysGlyIleLeuHisAspIleGluGlyIleArgGlyPro 213
DB 739 GAAAGATTACGTCGCTGCTGGCGCTTTTACATGATGATAGAACATCCCATTTGAAACA 798
QY 214 LysCysProLeuGlyValAspIleSerLysGlnValGlyLysIleAlaIleValPro 233
DB 799 AATATGCTCTTGGAGTACATATCCAGAAAGTTGCAAGAGCTGCAATAAAGTACCG 858
QY 234 GlnLeuLysMetGluIle 239
DB 859 CGATTAAATATGCACTT 876

```

RESULT 15
 BX930694 860 bp mRNA linear VRT 30-MAR-2004
 LOCUS Gallus gallus finished cDNA, clone CHEST967014.
 DEFINITION BX930694
 ACCESSION BX930694
 VERSION BX930694.2 GI:46016745
 KEYWORDS
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 860)
 Boardman, P.E., Bonfield, J.K., Brown, W.R.A., Carder, C., Chalk, S.E.,
 Croning, M.D.R., Davies, R.M., Francis, M.D., Grafham, D.V.,
 Hubbard, S.J., Humphrey, S.J., Hunt, P.J., Muddison, M., McLaren, S.R.,
 Niblett, D., Overton, I.M., Rogers, J., Scott, C.E., Taylor, R.G.,
 Tickle, C. and Wilson, S.A.
 Direct Submission
 Submitted (29-MAR-2004) Sanger Institute, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: chicken@hms.wmst.ac.uk

COMMENT

On Apr 1, 2004 this sequence version replaced gi:41631222.
 BBSRC/Dundee/Nottingham/Sanger/Sheffield/DMIST Gallus gallus cDNA
 sequencing project.
 This sequence is from the
 BBSRC/Dundee/Nottingham/Sanger/Sheffield/DMIST cDNA collection,
 from a library constructed by Elizabeth Bosch. cDNA was prepared
 from RNA extracted from limbs, normalised, and poly A-tailed.
 EcoRI-NciI cut cDNA was then ligated into the vector. Vector:
 pBlueScript II KS(+); Site_1: EcoRI; Site_2: NciI Host: Escherichia
 coli DH10B.

FEATURES

source
 1..860
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, HiseX"
 /db_xref="taxon:9031"
 /clone="CHEST967014"
 /clone_1ib="CSBCHN59"
 /dev_stage="stage 36"

ORIGIN

Alignment Scores:

Pred. No.: 7,866-71 Length: 860
 Score: 809.00 Matches: 151
 Percent Similarity: 79.66% Conservative: 37
 Best Local Similarity: 63.98% Mismatches: 48
 Query Match: 65.24% Indels: 0
 DB: 5 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x BX930694 (1-860)

```

QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspPheAla 20
DB 94 ATCTTGGACAGTACGACGAGAGGATTTCCTTCTGCTGACGATCATATGATATA 153
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 154 GCACCAAGTGCACATGCTGATTAAGTAGCAAGAGGCTCTTTAGTGAACACACCGGAG 213
QY 41 CysSerThrMetSerGlyLysValAlaIleGluHisLeuAlaLysGlnGlyAsnArgPhe 60
DB 214 TGCACAGCATGCGCGGAGGAAAGCAATAGGACCTGCTCAAAACCGAGACTGGCAA 273
QY 61 HisPheAspIleLysProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGly 80
DB 274 CAGTACACTTTCAGACTTCCCATGCAACAGATGTAATGATTTTCTTCTCCGGA 333
QY 81 LeuGlnHisValThrAspLysIleIleMetLysGlnGlyLysGlnGlyIleGluLys 100
DB 334 CTTCAAGAGCTTGTCAACAAAGCCATTCCTTCAAGAAAGAAAGAAAGATTTCAAGAA 393
QY 101 GlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCys 120
DB 394 GGGGAAATCCCTGCTCCTGTTAAGACATGCTCTCTGACACAGCATAGTGGCT 453
QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
DB 454 CATTTATCCAGCGGACACACCGACCATGCTCTTCTGCAATGAAAACAGATATTATTA 513
QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyValAlaAsnAsnPheThrIleArgArg 160
DB 514 CCAAAAAGTGAAGTCTGTTGTATCAAGAGAGTGAAGTAAATATCATATAGAAATA 573

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 16, 2005, 12:03:59 ; Search time 345.511 Seconds
(without alignments)
4094.859 Million cell updates/sec

Title: US-10-649-273-2_COPY_176_414
Perfect score: 1240
Sequence: 1 LIALVQGVSDFLIKGSLDI.....DISKEVGEASIKVPOLNMEI 239

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4390206 seque, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+g2n.model -DEV=xlp
-Q=/cgm2.1/USPTO.spool.p/US10649273/runat_15062005_111415_6022/app_query.fasta_1.1429
-DB=N Geneseq_16Dec04 -OPMT=fastcap -SUFFIX=g2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MTN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10649273_@CGN_1.1_1063_@runat_15062005_111415_6022 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:.*
1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002as:.*
7: geneseqn2002bs:.*
8: geneseqn2003as:.*
9: geneseqn2003bs:.*
10: geneseqn2003cs:.*
11: geneseqn2003ds:.*
12: geneseqn2004as:.*
13: geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1240	100.0	1416	8	ABX70950 Novel hum
2	1240	100.0	1526	6	ABX76639 DNA encod
3	1240	100.0	2058	6	ABX93268 Human O-g
4	1240	100.0	2197	6	ABX76635 DNA encod
5	1240	100.0	2572	8	ABX73207 Human pro

6	1203	97.0	1820	6	AAD46856 Human gty
7	1203	97.0	1820	6	ACA60887 Human cDN
8	1203	97.0	1821	10	ABX57020 DNA encod
9	1059	85.4	2208	10	ADAS2832 Human cod
10	1059	85.4	2890	12	ADQ24627 Human sof
11	995.5	80.3	3358	10	ADBE1345 Human dia
12	870	70.2	1572	6	ABQ75508 Murine si
13	599	48.3	2734	5	AA84622 DNA encod
14	468	37.7	371	12	ADL86725 DNA up-re
15	468	37.7	371	12	ADL86726 DNA up-re
16	332.5	26.8	1601	4	ABL24633 Drosophil
17	332.5	26.8	3656	4	ABL24632 Drosophil
18	332	26.0	1557	3	AA38454 Arabidops
19	257	20.7	1146	8	ACA26804 Prokaryot
20	248.5	20.0	1000	6	ABX37804 DNA seque
21	248.5	20.0	1000	6	ABX37804 DNA seque
22	248.5	20.0	1044	8	ACA39102 Prokaryot
23	248.5	20.0	94750	4	AA28551 Genomic f
24	247	19.9	936	8	ACA20445 Prokaryot
25	247	19.9	1053	12	ADL03120 DNA encod
26	241	19.4	1032	8	ACA43173 Prokaryot
27	239	19.3	4360	6	AD48239 Ehrlichia
28	235.5	19.0	1092	6	ABQ90383 M. capaul
29	234	18.9	600	3	AA400934 Human col
30	232	18.7	676	4	AA408019 Human cDN
31	232	18.7	1385	4	AA415110 Human cDN
32	231	18.6	1026	4	AA554064 Pseudomon
33	231	18.6	1026	8	ACA42146 Prokaryot
34	231	18.6	1026	10	ADG73341 P aerugin
35	231	18.6	1059	11	ABD02280 Pseudomon
36	231	18.6	1026	10	ABD02197 P aerugin
37	229	18.5	1206	10	ADG73343 P aerugin
38	224	18.1	1029	4	AA553309 Haemophil
39	224	18.1	1029	4	ACA34150 Prokaryot
40	224	18.1	11000	2	AA742063 O5 Continuation (6 of
41	217	17.5	9967	13	ADT05493 Haemophil
42	217	17.5	85814	13	ADT05493 Haemophil
43	214	17.3	1044	4	AB108591 Drosophil
44	213	17.2	1014	4	AA556045 Salmonell
45	213	17.2	1014	8	ACA51431 Prokaryot

ALIGNMENTS

RESULT 1	ABX70950	standard; CDNA; 1416 BP.
ID	ABX70950	
XX	AC	ABX70950;
XX	DT	05-MAR-2003 (first entry)
XX	DE	Novel human CDNA sequence #175.
XX	KW	Human; gene; ss; nervous system disorder; peripheral neuropathy;
XX	KW	Huntington's disease; amyotrophic lateral sclerosis; haemophilia;
XX	KW	neurodegenerative disease; Parkinson's disease; Alzheimer's disease;
XX	KW	autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;
XX	KW	insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;
XX	KW	ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;
XX	KW	fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;
XX	KW	coagulation disorder; cancer; tumour; inflammatory disease; septic shock;
XX	KW	Crohn's disease; anaphylaxis; proliferation; chemotactic;
XX	KW	differentiation; stem cell growth factor; hematopoiesis; chemokine;
XX	KW	haemostatic; antiinflammatory; expressed sequence tag; EST.
OS	XX	Homo sapiens.
XX	PN	W0200281731-A2.
XX	PD	17-OCT-2002.
XX	PF	29-JAN-2002; 2002WO-US001222.

XX 30-JAN-2001; 2001US-00774528.
XX (HYSE-) HYSEO INC.
XX (GOOD/) GOODRICH R. W.
XX Tang TY, Liu C, Zhou P, Aundi V, Zhang J, Zhao QA, Ren F;
XX Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;
XX MPI; 2003-058563/05.
XX
XX Novel polypeptide useful for treating neurodegenerative diseases, myeloid
XX or lymphoid cell disorders, bone disorders, mechanical and traumatic
XX disorders, coagulation disorders, and inflammatory diseases.
XX
XX Claim 1; Page; 612pp; English.
XX
XX This invention relates to the cDNA sequences encoding an isolated novel
XX human polypeptide. The protein encoded by the nucleic acid of the
XX invention is useful for treating central and peripheral nervous system
XX diseases (e.g. peripheral neuropathy, Huntington's disease, amyotrophic
XX lateral sclerosis); neurodegenerative diseases (e.g. Parkinson's disease,
XX Alzheimer's disease); autoimmune disease (e.g. systemic lupus
XX erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus)
XX ; myeloid or lymphoid cell disorders (e.g. anaemia and thrombocytopenia)
XX ; wounds, ulcers, burns; bone disorders (e.g. osteoporosis,
XX osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head
XX trauma); lung or liver fibrosis; reperfusion injury in various tissues;
XX bacterial, viral or fungal infections; allergic conditions such as
XX allergic rhinitis, asthma; coagulation disorders (e.g. haemophilia);
XX cancer and tumours; and inflammatory diseases (e.g. septic shock, Crohn's
XX disease, anaphylaxis). The protein may be used to inhibit the growth,
XX infection or function of infectious agents such as bacteria, fungi,
XX viruses, or to effect bodily characteristics, biorhythms or circadian
XX cycles of rhythms. The protein may also have
XX proliferation/differentiation, stem cell growth factor, haematopoiesis
XX regulation, immune stimulation or suppressing, chemotactic/chemokinetic,
XX haemostatic and thrombolytic, receptor/ligand, and anti-inflammatory
XX activities. The cDNA sequences of the invention are useful for expressing
XX recombinant protein for analysis. The present sequence represents a novel
XX human cDNA sequence of the invention, this sequence is an expressed
XX sequence tag (EST) and was identified using subtractive hybridisation
XX
SQ Sequence 1416 BP; 441 A; 280 C; 273 G; 422 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3,21e-132 Length: 1416
Score: 1240.00 Matches: 239
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0
US-10-649-273-2_COPY_176_414 (1-239) x ABX70950 (1-1416)
QY 1 LeuLeuAlaIeuValGlnGlyValSerAspPheIeuIeuGlyIysSerIeuAspIle 20
DB 586 CTTGTGGCATTATTCAGAGAGTTTCAGATTTCTGCTTCTGGAAAGCTCTTGGACACTA 645
QY 21 AlaProGlyAspMetLeuAspIysValAlaIArgArgIeuSerIeuIleIysHisProGlu 40
DB 646 GCACCAAGTCACATGCTTGCACAGGTGGCAAGAGACTTTCTTAAATAAACATCCAGAG 705
QY 41 CysSerThrMetSerGlyGlyValAlaIleGluHisIeuAlaIysGlnGlyAsnArgPhe 60
DB 706 TGTCTCCACCATGAGTGGTGGAAAGCATAGACATTGGCCAAACAGAAATATGATTT 765
QY 61 HisPheAspIleIysProProIeuHisHisAlaIysAsnGlyAspPheSerPheThrGly 80
DB 766 CATTTGACATCAAACTCCCTTGCATCATGCTAAAAATGATGATTTTCTTTTACTGGA 825
QY 81 LeuGlnHisValThrAspIysIleIleMetIysGlnGlyIysGlnGlnGlyIleGluIys 100

DB 826 CTTCAACACGTTACTGATTAATAATATGAAAAAGAAAAAGAGAGATTATTGAGAG 885
QY 101 GlyGlnIleIeuSerSerAlaAlaAspIleAlaIleThrValGlnHisThrMetAlaCys 120
DB 886 GGGCAATCTCTGCTTTCAGCAGCAGCATTCCTCCACAGTACAGCACAATGAGCATGT 945
QY 121 HisIeuValIysArgThrHisIArgAlaIleIeuPheCysIysGlnArgAspIeuPro 140
DB 946 CATCTTGAAAAAGAACACATCGGGCTATTCTGTTTGAAGCAGAGACCTGTACTCT 1005
QY 141 GlnAsnAsnAlaValIeuValAlaSerGlyGlyValAlaSerAsnPheThrIleArgArg 160
DB 1006 CAAATATATCAGATGATGCTGGTGCATCGGTGCTCCAACTAATCTTATATTCGACGA 1065
QY 161 AlaIeuGlnIleIeuThrAsnAlaThrGlnCysThrIeuIeuCysProProArgIeu 180
DB 1066 GCTTGGAAATTTTAAACAAACGACACAGTGCATTTGTGTCTCTCCACAGACTA 1125
QY 181 CysThrAspAsnGlyIleMetIleAlaIleArgAsnGlyIleGluArgIeuAlaGlyIeu 200
DB 1126 TGCACTGATATGACATTATGATTCATGCAAGATGCTATGAAAGACTACGTGCTTG 1185
QY 201 GlyIleIeuHisAspIleGlnGlyIleArgGlyIleGluProIysCysProIeuGlyValAsp 220
DB 1186 GGCATTTTACATGACATAGAAAGCATCCGCTATTAACCAAAATGTCTCTTGGAGTAGAC 1245
QY 221 IleSerIysGlnValGlyValAlaSerIleIysValProGlnIeuIysMetGluIle 239
DB 1246 ATATCAAAAGAGATGGAGAGCTTCATTAAGTACACATTAATAATGAGATA 1302
RESULT 2
ABST76639
ID ABST76639 standard; DNA; 1526 BP.
AC
XX ABS76639;
XX
XX 11-DEC-2002 (first entry)
XX
DE DNA encoding novel human metalloprotease MP1 fragment #1.
XX
XX Metalloprotease; MP-1; immune disorder; glutamate transport; cancer;
XX motor neuron disorder; amyotrophic lateral sclerosis; ALS; diabetes;
XX reproductive disorder; Klinefelter's syndrome; germinal cell aplasia;
XX genital wart; metabolic disorder; premature puberty; Kallman syndrome;
XX Cushing's syndrome; neurodegenerative disease; Alzheimer's disease;
XX Parkinson's disease; Huntington's disease; Tourette syndrome; sepsis;
XX liver disease; renal disease; immune disorder; rheumatoid arthritis;
XX acquired immunodeficiency syndrome; AIDS; pulmonary disease; pneumonia;
XX emphysema; cystic fibrosis; vascular disorder; inflammatory disorder;
XX neurological disorder; gene; ds.
OS Homo sapiens.
XX
XX MO200272751-A2.
XX
XX 19-SEP-2002.
XX
XX 05-FEB-2002; 2002W0-US003353.
XX
XX 05-FEB-2001; 2001US-0266518P.
XX 10-APR-2001; 2001US-0282814P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Chen J, Feder J, Nelson TC, Duclos F, Krystek S;
XX
XX MPI; 2002-723329/78.
XX P-PSDB; ABG96487.
XX
XX New isolated nucleic acid encoding MP-1 protein, useful for preventing,
XX treating, or ameliorating diseases associated with aberrant
XX metalloproteinase activity, e.g. immune, metabolic, inflammatory and
XX neurological disorders.

XX Disclosure; Page 462-463; 473pp; English.

CC The invention describes an isolated nucleic acid molecule (I) encoding a
CC metalloprotease (MP-1). (I) is useful for preventing, treating, or
CC ameliorating a medical condition, particularly an immune disorder, an
CC aberrant glutamate transport or motor neuron disorder, such as
CC amyotrophic lateral sclerosis (ALS), its juvenile form or an ALS-like
CC condition. The composition and methods are also useful for diagnosing,
CC prognosticating, treating, ameliorating and/or treating disorders
CC associated with MP-1 activity, e.g. diabetes, cancer, reproductive
CC disorders (e.g. Klinefelter's syndrome, genital warts, or germinal cell
CC aplasia), metabolic disorders (e.g. premature puberty, Kallman syndrome,
CC or Cushing's syndrome), neurodegenerative diseases (Alzheimer's disease,
CC Parkinson's disease, Huntington's disease or Tourette syndrome), liver
CC and renal diseases and immune disorders (e.g. AIDS, rheumatoid arthritis
CC or sepsis), pulmonary diseases (e.g. pneumonia, emphysema or cystic
CC fibrosis) and vascular, inflammatory and neurological disorders (e.g.
CC Alzheimer's disease or Parkinson's disease). This sequence represents a
CC metalloprotease MPI polynucleotide

XX Sequence 1526 BP; 484 A; 297 C; 289 G; 456 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3,57e-132	Length:	1526
Score:	1240.00	Matches:	239
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-649-273-2_COPY_176_414 (1-239) x ABA93268 (1-1526)

```
Qy 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
Db CTTGTGGCATTAGTTCAGAGAGTTTCAGATTCTGCTTCTTGAAAGCTTTGGACATA 144
Qy 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
Db GCACCAAGTGACATGCTTGAACAAGTGGCAAGAACCTTCTTAAATAAACATCCAGAG 204
Qy 41 CysSerThrMetSerGlyLysValAlaIleGlnHisLeuAlaLysGlnGlyAsnArgPhe 60
Db TGGTCCACCAAGTGAGTGGGAAAGCCATACACATTGGCCAAACAGAAATAGATT 264
Qy 61 HisPheAspIleLysProLeuHisHisAlaLysAsnCysAspPheSerPheThrGly 80
Db CATTGGACATCAAACTCCCTTCATCATCTAAATAATGATTTCTTTACTGGA 324
Qy 81 LeuGlnHisValThrAspLysIleIleMetLysLysGluGluGlyIleGluLys 100
Db CTTCAACACGTTACTGATTAATAATATGAAAAAGGAAAAAGGAGATTTGAGAG 384
Qy 101 GlnGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCys 120
Db GGGCAATATCTGCTTTCAGACAGACATGCTCCACAGTACAGACACAAATGGCATGT 444
Qy 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
Db CATCTTGAGAAAGACACATCGGGCTATTTCTGTTTGTAAAGCAGAGACTTTGACT 504
Qy 141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAspPheTyrIleArgArg 160
Db CAAATAATATGACAGTACGTTGATCGTGTGCGCAATCTTCTATATATCGCAGA 564
Qy 161 AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLeu 180
Db GCTCTGAAATTTTAAACAAACGCAACAGTGCATTTGTGTGTCTCCCTCCACAGACTA 624
Qy 181 CysThrAspArgGlyIleMetIleAlaThrAsnGlyIleGluArgLeuArgAlaGlyLeu 200
Db TGACATGATATATGACATTATGATTCAGTGAATGATTTGAAAGCTACGTCGCTTG 684
```

```
Qy 201 GlyIleLeuHisAspIleGluGlyIleArgTyrGluProLysCysProLeuGlyValAsp 220
Db 685 GGCATTTTACATGACATAGAGGATCGCTATGAAACAAATATGCTCTTGGAGTAGAC 744
Qy 221 IleSerLysGluValGlnGlyAlaSerIleLysValProGlnLeuLysMetGluIle 239
Db 745 ATATCAAAAAGAGTTGGAGAGCTTCCATTAAGTACACAACTTAAATATGAGATA 801
```

RESULT 3

ABA93268 standard; cDNA; 2058 BP.

AC ABA93268;

DT 19-APR-2002 (first entry)

DE Human O-sialoglycoproteinase-like protein encoding cDNA SEQ ID NO:1.

XX Human; O-sialoglycoproteinase-like protein; OSGPLP; enzyme; gene; ss.

XX Homo sapiens.

OS Homo sapiens.

FT Key Location/Qualifiers

FT CDS 110..1354

FT FT /*tag= a

FT FT /product= "O-sialoglycoproteinase-like protein"

PD CNJ138550-A.

PD 24-OCT-2001.

PP 19-APR-2000; 2000CN-00106834.

PP 19-APR-2000; 2000CN-00106834.

PR (SHAN-) SHANGHAI SHENGYUAN GENE DEV CO LTD.

PA Mao Y, Xie Y,

PI WPI: 2002-115090/16.

DR P-PSDB; ABB05481.

XX O-sialoglycoproteinase-like protein and encoding polynucleotide, useful

PT for diagnosing, preventing and treating related diseases.

XX Claim 5; Page 29-30 (Disclosure); 38pp; Chinese.

PS The present sequence encodes human O-sialoglycoproteinase-like protein

CC (OSGPLP). The present invention also describes: (1) the preparation of

CC the OSGPLP protein; (2) applying the OSGPLP protein in diagnosis; (3) the

CC prevention and/or treatment of related diseases; (4) utilizing the OSGPLP

CC an antibody against the OSGPLP protein; and (5) the use of the OSGPLP

CC polynucleotide sequences, proteins, agonists, excitors, inhibitors

CC and antibodies in treating diseases related to the abnormal OSGPLP gene

XX and in preparing the medicine composite for the treatment

XX Sequence 2058 BP; 637 A; 400 C; 410 G; 611 T; 0 U; 0 Other;

US-10-649-273-2_COPY_176_414 (1-239) x ABA93268 (1-2058)

```
Qy 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
Db CTTGTGGCATTAGTTCAGAGAGTTTCAGATTCTGCTTCTTGAAAGCTTTGGACATA 694
```

QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 695 GCAACGAGTGAAGCTTGAAGAGTGGCAAGAGCTTTCTTTAATATAAATCCAGAG 754
QY 41 CysSerThrMetSerGlyGlyValAlaIleGluHisLeuAlaLysGlnGlyAsnArgPhe 60
DB 755 TGTCTCCACATGATGATGGTGGAAAGCCATAGAAATTGGCCAAAGCAAGAAATAGATTT 814
QY 61 HisPheAspIleLysProProLeuHisHisAlaLysAsnGlySAspPheSerPheThrGly 80
DB 815 CATTGTGACATCAACCTCCCTTGACATCATGCTTAAATATGATTTTCTTTTACTGGA 874
QY 81 LeuGlnHisValThrAspLysIleIleMetLysLysGluGlyGluGlyIleGluLys 100
DB 875 CTTCAACAGCTTACTGATATAATATATGAAAAAGAAAAAGAGAAAGATTTGAGAG 934
QY 101 GlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCys 120
DB 935 GGGCAATCTCTGCTTCACAGCAGACATTTGCTGCCACAGTACAGCACATGGCATGT 994
QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
DB 995 CATCTTGATAAAGAACACATCGCGCTATTCTGTTTGTAAAGCAGAGACTGTCTTAACT 1054
QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPheTyrIleArgArg 160
DB 1055 CAAATAAATGCACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1114
QY 161 AlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLeu 180
DB 1115 GCTCTGAAATTTTAAACAAACGCAACACAGTGCACCTTTGTGTGTCCTCCAGACTA 1174
QY 181 CysThrAspAsnGlyIleMetIleAlaThrAsnGlyIleGluArgLeuAlaGlyLeu 200
DB 1175 TGCACGATTAATGATGATTAATGATGATGATTAATGATTAATGATGATGATGATG 1234
QY 201 GlyIleLeuHisAspIleGluGlyIleArgTyrGluProLysCysProLeuGlyValAsp 220
DB 1235 GGCATTTTACATACATAGAGGATCCGCTATGAAACCAAAATGTCCTTTGGAGTAGAC 1294
QY 221 IleSerLysGluValGlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 239
DB 1295 ATATCAAAAGAAAGTTGGAGAGAGCTTCATTAAGTATCACACATTAATAAATGAGATA 1351

RESULT 4

ABST76635
ID ABST76635 standard; DNA; 2197 BP.

XX AC ABST76635;

XX DT 11-DEC-2002 (first entry)

XX DE DNA encoding novel human metalloprotease MP1.

KW Metalloprotease; MP-1; immune disorder; glutamate transport; cancer;
motor neuron disorder; amyotrophic lateral sclerosis; ALS; diabetes;
reproductive disorder; Klinefelter's syndrome; germinal cell aplasia;
genital wart; metabolic disorder; premature puberty; Kallman syndrome;
Cushing's syndrome; neurodegenerative disease; Alzheimer's disease;
Parkinson's disease; Huntington's disease; Tourette syndrome; sepsis;
liver disease; renal disease; immune disorder; rheumatoid arthritis;
acquired immunodeficiency syndrome; AIDS; pulmonary disease; pneumonia;
emphysema; cystic fibrosis; vascular disorder; inflammatory disorder;
neurological disorder; gene; ds.

XX OS Homo sapiens.

XX PN WO200272751-A2.

XX PD 19-SEP-2002.

XX PF 05-FEB-2002; 2002WO-US003353.

PR 05-FEB-2001; 2001US-0266518P.
PR 10-APR-2001; 2001US-0282814P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
PI Chen J, Feder J, Nelson TC, Duclos F, Krystek S;
XX
DR MPI; 2002-723329/78.
DR P-PSDB; ABG96478.
XX
PT New isolated nucleic acid encoding MP-1 protein, useful for preventing,
PT treating, or ameliorating diseases associated with aberrant
PT metalloproteinase activity, e.g. immune, metabolic, inflammatory and
PT neurological disorders.

PS Claim 1; Fig 1A-C; 473pp; English.

CC The invention describes an isolated nucleic acid molecule (i) encoding a
CC metalloprotease (MP-1). (i) is useful for preventing, treating, or
CC ameliorating a medical condition, particularly an immune disorder, an
CC aberrant glutamate transport or motor neuron disorder, such as
CC amyotrophic lateral sclerosis (ALS), its juvenile form or an ALS-like
CC condition. The compositions and methods are also useful for diagnosing,
CC prognosticating, treating, ameliorating and/or treating disorders
CC associated with MP-1 activity, e.g. diabetes, cancer, reproductive
CC disorders (e.g. Klinefelter's syndrome, genital warts, or germinal cell
CC aplasia), metabolic disorders (e.g. premature puberty, Kallman syndrome,
CC or Cushing's syndrome), neurodegenerative diseases (Alzheimer's disease,
CC Parkinson's disease, Huntington's disease or Tourette syndrome), liver
CC and renal diseases and immune disorders (e.g. AIDS, rheumatoid arthritis
CC or sepsis), pulmonary diseases (e.g. pneumonia, emphysema or cystic
CC fibrosis) and vascular, inflammatory and neurological disorders (e.g.
CC Alzheimer's disease or Parkinson's disease). This sequence represents a
CC metalloprotease MP1 polynucleotide

SQ Sequence 2197 BP; 681 A; 441 C; 439 G; 636 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	2197
Score:	5.98e-132	
Percent Similarity:	1240.00	Matches: 239
Best Local Similarity:	100.00%	Conservative: 0
Query Match:	100.00%	Mismatch: 0
DB:	6	Indels: 0
		Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x ABST76635 (1-2197)

QY 1 LeuLeuAlaLeuValGlnGlyValIleSerAspPheLeuLeuGlyLysSerLeuAspIle 20
DB 756 CTGTGGCATTAAGTTCAGAGAGTTTCAGATTTCTGCTTCTTGAAAGTCTTTGAGACATA 815
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 816 GCACGAGTGAAGCTTGAAGAGTGGCAAGAGACTTTCTTTAATATAAATCCAGAG 875
QY 41 CysSerThrMetSerGlyGlyValAlaIleGluHisLeuAlaLysGlnGlyAsnArgPhe 60
DB 876 TGTCTCCACATGATGATGGTGGAAAGCCATAGAAATTTGGCCAAAGCAAGAAATAGATTT 935
QY 61 HisPheAspIleLysProProLeuHisHisAlaLysAsnGlySAspPheSerPheThrGly 80
DB 936 CATTGTGACATCAACCTCCCTTGACATCATGCTTAAATATGATTTTCTTTTACTGGA 995
QY 81 LeuGlnHisValThrAspLysIleIleMetLysLysGluGlyGluGlyIleGluLys 100
DB 996 CTTCAACAGCTTACTGATATAATATATGAAAAAGAAAAAGAGAAAGTATTGAGAG 1055
QY 101 GlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCys 120
DB 1056 GGGCAATCTCTGCTTCACAGCAGACATTTGCTCCACAGTACAGCACATGGCATGT 1115
QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140

Db	1116	CATCTTGGAAGAAGAACACATCGGCGTATCTGTTTGTAAGCAGAGACTGTACTT	1175
Qy	141	Glma5n5n5n1a1aVal1euVal1Ala5erG1yG1yVal1a1a1a5erA5nPhaTtY11eArGr	160
Db	1176	CAAAATATATGACATACGTGGTGCATCTGGTGGTGGCAAGTAACTTCTATATCCGACGA	1235
Qy	161	Ala1euG1u11e1euThrA5nAlaThrG1ncY5erTh1euLeuCy5eProP5roPhrGheu	180
Db	1236	GCTCTGGAAATTTTACACAAAGCCACACACAGTGCACCTTGTGTGCTCTCCACGACTA	1295
Qy	181	Cy5ThrA5pA5nG1y11eMeT11eAla1ArpA5nG1y11eG1uArG1euArGAlaG1y1eu	200
Db	1296	TGCACGTAGTAAATGGCATATATGATTCGATGCAGATG5TATGAAAGACTAGTGGCTTG	1355
Qy	201	G1y11e1euH1a5pH11eG1uG1y11eArGrG1y11eArGrG1y11eArGrG1y11eArGr	220
Db	1356	GGCATTTTACATGACATAGAAAGGCATCCGCTATGAAACCAAAATGCTCTTGGAGTACAC	1415
Qy	221	11e5er1y5G1uVal1G1yG1uAla5er11e1y5Val1ProG1nc1eu11e1y5MeT11e	239
Db	1416	ATATCAAAAGAAAGTTTGAAAGCTTCCATTAAGTACCAACATTTAAATAGAGATA	1472
RESULT 5			
ABT23207	ID	ABT23207	standard; DNA; 2572 BP.
XX	ABT23207;		
XX	01-MAY-2003	(first entry)	
XX	Human protein modification + maintenance molecule DNA SEQ ID NO 36.		
DE	Cytostatic; antiarteriosclerotic; anti-HIV; antiallergic; anticonvulsant;		
XX	cerebroprotective; antiparkinsonian; nootropic; antiinflammatory;		
KW	antitumor; hepatotropic; gynaecological; antibacterial; virucide;		
KM	protozoacide; antiparasitic; cell proliferative disease; PMOD;		
KM	protein modification and maintenance molecule; immunogenic fragment;		
KW	cancer; autoimmunity; inflammatory disease; neurological disorder;		
KW	gastrointestinal; developmental; vesicle trafficking disorder; infection;		
KW	protein-protein interaction; drug-target interaction;		
XX	gene expression profile; human; gene; ds.		
OS	Homo sapiens.		
XX	WO2003000844-A2.		
XX	03-JAN-2003.		
PD	18-JUN-2002; 2002WO-US019360.		
PF	22-JUN-2001; 2001US-0300508P.		
XX	06-JUN-2001; 2001US-0303445P.		
PR	13-JUN-2001; 2001US-0305405P.		
PR	09-AUG-2001; 2001US-0311442P.		
PR	24-AUG-2001; 2001US-0314821P.		
PR	29-AUG-2001; 2001US-0315929P.		
PR	03-MAY-2002; 2002US-0378205P.		
XX	(INCY-) INCYTE GENOMICS INC.		
PA	Gandhi AR, Kable AE, Swarnakar A, Hafalia AJA, Tran B, Duggan BM,		
XX	Warren BA, Ison CH, Honchell CD, Nguyen DB, Lee EA, Yue H,		
PI	Foreythe BJ, Barroso I, Ramkumar J, Griffin JA, Li JX, Yang J,		
PI	Thangavelu K, Gietzen KJ, Ding L, Baughn MR, Borowsky MU, Yao MG,		
PI	Walia NK, Mason PM, Gururajan R, Lee S, Becha SD, Lee SY, Tran UK,		
PI	Elliot VS, Luo W, Sprague WM, Tang YT, Lu Y, Zedajadian Y,		
XX	WPI; 2003-184039/18.		
DR	P-PSDB; AB026654.		
XX	New isolated human PMOD polypeptide and polynucleotide, useful for		
PT	diagnosing, treating and preventing diseases or conditions associated		
PT	with the aberrant PMOD expression,e.g. cancer, AIDS, atherosclerosis and		

PX infections.

PS Claim 91; Page 211; 225pp; English.

XX

The invention relates to an isolated polypeptide comprising: any of 28 sequences of 48-1256 amino acids; a natural amino acid sequence at least 90% identical to the 28 amino acid sequences, 94% identical to a sequence CC of 703 or 267 amino acids, 96% identical to a sequence of 414 amino acids, or 97% identical to a sequence of 242 amino acids, all given in the specification; or a biologically active or immunogenic fragment of the isolated polypeptide. The polypeptides and polynucleotides are useful in diagnosing, treating and preventing diseases or conditions associated with the decreased expression of protein modification and maintenance molecules (PMOD), such as cell proliferative diseases (e.g. cancer, atherosclerosis), autoimmune/inflammatory diseases (e.g. AIDS, allergies), neurological disorders (e.g. stroke, Parkinson's disease, epilepsy), gastrointestinal (e.g. ulcer, cirrhosis), reproductive (e.g. endometriosis) developmental, vesicle trafficking disorder, and infection (e.g. bacterial, viral, parasitic, protozoal). These are also useful in assessing the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of PMOD. The PMOD or its fragments are useful in screening compounds for effectiveness as agonist or antagonist of the polypeptides, or in altering the expression of the target polynucleotide and compounds that specifically bind to or modulate the activity of the polypeptide. The microarray is useful in monitoring or measuring protein-protein interactions, drug-target interactions, and gene expression profiles. This polynucleotide sequence represents the DNA encoding a human PMOD protein of the invention

SQ Sequence 2572 BP; 780 A; 489 C; 509 G; 794 T; 0 U; 0 Other;

Alignment Scores:

Folded NO. :	Length:	Score:	Matches:	Conservative:	Mismatches:	Gaps:
US-10-649-273_2_COPY_176_414 (1-239)	x ABR23207 (1-2572)	7.47e-132	2572	239	0	0
OY 1 LeuLeuAlaValGlnGlyVaiSerAspPheIleuLeuLysSeriLeuAspIle	20	1240.00	239	0	0	0
Dd 669 CTGTGGCATTAAGTTCACAGAAGCTTTCCATTTCTTGGAAGACTTTTAGCAATA	728	Percent Similarity: 100.00%	Indels: 0			
OY 21 AlaProGlyAseMetLeuAspLysVaIAlaArgTrgLeuSeriLeuHisProGlu	40	Query Match: 100.00%	Indels: 0			
Dd 729 GCACCAGGTGCATCGCTTGACAAGGTGCCAAAGAACATTCTTTAATTAACATCCAGAG	788					
OY 41 CysSerThrMetSerGlyGlyLysAlaIlegluniSteuAlalyGngilyAsnArghPhe	60					
Dd 789 TGCTTCACCAAGATGTGTGGGAAAAGCCTACAAACATTGGCCAACAAAGAAAAAGATTT	848					
OY 61 HisPheAspIleLysProProLeuHisHiaIalaLysAenCySaAPRheserPhethrGly	80					
Dd 849 CATTTTGACATCAAACCTCCCCTGCATCATCTAAAAATGTGATTTTTCTTTAATCGGA	908					
OY 81 LeuGlunhisValThraAspLysileil emetLysHySGiulySGiugLyylEGLiudyS	100					
Dd 909 CTTCAACACGTTACTGANTBAATBATAATGAAAAAGAAAAAGAAAGGATATTCAGAG	968					
OY 101 GlyGlnIleLeuSerSerAlaAlaAspIleAlaAlatrValGlnhiFthrMeCaLYCs	120					
Dd 969 GGCGAAATCCGTCTTCAGCACGACACATTCCTGCCACAGTACAGCACACATGGCANGT	1028					
OY 121 HiseuVallysaaqutrrhiaargAlaleuephecyelysginaraasphleuleupro	140					
Dd 1029 CATCTTGTGAAAAGAACACATCGGGCTAATCTGTGTTTGTAAGCAGAGAGACTTGTA	1088					
OY 141 GlusnaSenAlavalLeuvalAlaserGIyGVaialaSerasnPhetyril eargrar	160					
Dd 1089 CAANATATAGCAGTACGTGTGTCATCTGTGTGTGTGCGAAGTAACTTCTATATCCGCA	1148					

QY	161	AlaLeuGluIleLeuThrAsnAlaThrGlnCysTrpLeuLeuCysProProPargLeu	180
Db	1149	GCTCTGGAAATTTTAACAAACCAACACAGTGCACCTTGTCTTGTCTCTCCAGACTA	1208
QY	161	CysThrAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeu	200
Db	1209	TGCACATGATTAATGCGATTATATGATTCATGATGAAATGATTAATGAAGACTACGTCGTGGCTTG	1268
QY	201	GlyIleLeuHisAspIleGluGlyIleLeuArgTrpGluProGlyCysProLeuGlyValAsp	220
Db	1269	GGCATTTTAAACAGACATGAGAAAGCATCCGCTATGAAACCAAAATGCTCTTGAGGTAC	1328
QY	221	IleSerIysGluValGlyValAsnSerIleLysValProGlnLeuLysMetGlnIle	239
Db	1329	ATATCAAAAGAGCTTGAGAGACCTTCATTAAGTACCAATTAATAATGAGATA	1385
RESULT 6			
AAD46856			
ID	AAD46856	standard; cDNA; 1820 BP.	
XX			
AC	AAD46856;		
DT	27-JAN-2003	(first entry)	
XX			
DE	Human glycoprotease 28472 cDNA.		
XX			
KM	Human, adenosine deaminase; seven transmembrane domain receptor; cancer;		
KM	7TM; glycoprotease; immune disorder; Iga deficiency; allergy; arthritis;		
KM	rheumatoid arthritis; diabetes; atherosclerosis; cardiovascular disorder;		
KM	hypertension; ischaemic heart disease; obesity; myocardial infarction;		
KM	endothelial cell disorder; Grave's disease; psoriasis; brain disorder;		
KM	Parkinson's disease; Alzheimer's disease; hematopoietic disorder;		
KM	cerebral edema; metabolic disorder; liver disorder; platelet disorder;		
KM	chromosome mapping; tissue typing; gene therapy; neuroprotective;		
KM	cytostatic; anorectic; cardiant; haemostatic; gene; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	146..1390	
FT		/*tag = a	
FT		/product = "Human 28472 protein"	
FT		/note = "This region is specifically claimed as SEQ ID NO:	
FT		6 in claim 1 of the specification"	
PN	WO200274960-A2.		
XX			
PD	26-SEP-2002.		
XX			
PF	08-NOV-2001; 2001WO-US051427.		
XX			
PR	08-NOV-2000; 2000US-0246768P.		
PR	08-NOV-2000; 2000US-0246772P.		
PR	15-NOV-2000; 2000US-0249185P.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
PI	Leibyl KR, Kapeller-Libermann R, Glucksmann M;		
XX			
DR	WPI; 2002-759898/82.		
XX	P-PSDB; AAE29234.		
XX			
PT	New 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid molecules,		
PT	useful for diagnosing and treating cancer, immune, cardiovascular,		
PT	hematopoietic, brain, pain, metabolic, liver or platelet disorders, and		
PT	in pharmacogenomics.		
PS	Claim 1; Fig 8; 178pp; English.		
XX			
CC	The present invention relates to novel 38650, 28472, 5495, 65507, 81588		
CC	or 14354 nucleic acid molecules which encode adenosine deaminase, glyco-		
CC	protease or seven transmembrane domain (7TM) receptor family members.		
CC	Sequences of the invention are useful in diagnosing and treating cancer		

CC	or aberrant cellular proliferation and/or differentiation (e.g. colon or lung cancer), immune disorders (e.g. selective IgA deficiency, rheumatoid arthritis, diabetes, allergy), heart or cardiovascular disorders (e.g., hypercortisolemia, atherosclerosis, arrhythmias, ischemic heart disease, myocardial infarction, thrombosis), including endothelial cell disorders (e.g. psoriasis, Graves' disease), haematopoietic disorders, brain disorders (e.g. cerebral edema, Parkinson's or Alzheimer's disease), pain and metabolic disorders (e.g. obesity), liver disorders or platelet disorders. They are also useful in screening assays, predictive medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical trials and pharmacogenetics) and prophylactic and therapeutic methods. The nucleic acids may also be used in chromosome mapping, tissue typing and forensic biology and as surrogate markers. Sequences of the invention are also used in gene therapy. The present sequence is human glycoprotease 28472 cDNA
CC	XX
CC	Sequence 1820 BP; 543 A; 365 C; 393 G; 518 T; 0 U; 1 Other;
CC	Alignment Scores:
CC	Pred. No.: 8 23e-128 Length: 1820
CC	Score: 1203.00 Matches: 232
CC	Percent Similarity: 98.33% Conservative: 3
CC	Best Local Similarity: 97.07% Mismatches: 4
CC	Query Match: 97.02% Indels: 0
CC	DB: 6 Gaps: 0
CC	US-10-649-273-2_COPY_176_414 (1-239) x AAD46856 (1-1820)
CC	1 LeuLeuAlaLeuValGInGlyValSerAspPheLeuLeuLeuGlyLysSerLeuAspIle 20
CC	671 CTGTGGCATTAGTTCAGAGGTTTCAGATTTCCTGCTTCAGAAAGTCTTGACATA 730
CC	21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
CC	731 GCACCAAGGTGACATGCTTGCACCAAGGTGCACAAAGACTTCTTTAAATAAACATCCAGAG 790
CC	41 CysSerThrMetSerGlyLysValAlaIleGlnHisLeuAlaLysGlnGlyAsnArgPhe 60
CC	791 TGCTCCACCAAGATGGTGGGAAAGCCATAGAACATTGGCCAAACAGAAATAGAGATT 850
CC	61 HisPheAspIleLysProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGly 80
CC	851 CATTTTGACATCAACACTCCCTTGACATCATGCTCTAAAATTGTGATTTCTTTACTGGA 910
CC	81 LeuGlnHisValIleThrAspLysIleIleMetLysLysGlyLysGlnGlyIleGlnLys 100
CC	911 CTTCAACACGCTTACTGATTAATAATATGAAACACAGAAACACAGAGGATTTGAGAAG 970
CC	101 GlyGlnIleLeuSerSerSerAlaAlaAspIleAlaIleThrValGlnHisSerMetAlaCys 120
CC	971 GGGCAAAATCCGTCTTGACGACGACGACGACATTTGCGCACATGACGACACACATGCAATG 1030
CC	121 HisLeuValLysArgThrHisSerGlyAlaIleLeuPheCysLysGlnArgAspLeuLeuPro 140
CC	1031 CATCTTGTGAAGAAGAACATCGGGCTATTTCTGTTTGTAAAGCAGAGACTTGTACTT 1090
CC	141 GlnAsnAsnAlaValLeuValAlaSerGlyValAlaAserAsnPheTyrIleArgArg 160
CC	1091 CAATATATGACAGTACTGGTTGATCTGGTGTGTGTCGCAAGTAACTCTTATATCCGAGA 1150
CC	161 AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrIleLeuLeuCysProProProArgLeu 180
CC	1151 GCTCTGGAATTTTAAACAAACGCAACAGGACCTTGTGTGTCTCTCCACGACTA 1210
CC	181 CysThrAspAsnGlyIleMetIleAlaTyrAsnGlyIleGluArgLeuArgAlaGlyLeu 200
CC	1211 TGCACTGATATATGCAATTATGATTGATCGAATGTATTTGAAGAATCAATGCTGCGCTTG 1270
CC	201 GlyIleLeuHisAspIleGlnGlyIleArgTyrGlnProLysCysProLeuGlyValAsp 220
CC	1271 GGCAATTTTAACTACATAGAAAGGATCCGGATGATGAACCAAAATGTCTCTTGAGTGAAC 1330
CC	221 HisSerLysGlnValGlyGlnAlaSerTyrIleLysValProGlnLeuLysMetGlnIle 239

DB 1331 ATATCAAAAGAGTGGAGAGCTTCATTAAGTACACACATTAAAAATGAGATA 1387
RESULT 7
ACA60887
ID ACA60887 standard; cDNA, 1820 BP.
XX
AC ACA60887;
XX
DT 08-JUL-2003 (first entry)
XX
DE Human cDNA 28472 encoding a glycoprotease.
XX
KW Human; ss; gene; cancer; aberrant cellular proliferation;
KW differentiation; immune disorders; heart disorder; brain disorder;
KW cardiovascular disorder; endothelial cell disorder; pain disorder;
KW haematopoietic disorder; blood vessel disorder; metabolic disorder;
KW liver disorder; platelet disorder; glycoprotease.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 146..1390
FT /tag= a
FT /product= "glycoprotease"
FT /note= "this CDS is specifically claimed in claim 1"
XX
PN US2003009017-A1.
XX
XX 09-JAN-2003.
PD 08-NOV-2001; 2001US-00012140.
XX
PF 08-NOV-2000; 2000US-0246768P.
XX PR 08-NOV-2000; 2000US-0246772P.
XX PR 15-NOV-2000; 2000US-0249185P.
XX
PA (LEIB/) LEIBY R.
PA (KAPE/) KAPELIER-LIBERMANN R.
PA (GLUC/) GLUCKSMANN M A.
PI
PI Leiby KR, Kapellier-Libermann R, Glucksmann MA;
DR MPI; 2003-428888/40.
DR P-PDB; AB009569.
XX
PT New isolated 38650, 28472, 5495, 65507, 81588 and 14354 nucleic acid
PT molecules, useful for diagnosing, treating cancer, pain, or immune,
PT heart, endothelial cell, hematopoietic, blood vessel, brain, metabolic
PT and liver disorders.
XX
XX Claim 2; Fig 8; 90pp; English.
PS
XX The invention relates to an isolated 38650 (encoding adenosine
CC deaminase), 28472 (encoding a glycoprotease), 5495 (encoding a 7
CC transmembrane receptor), 65507, 81588 or 14354 nucleic acid molecule (or
CC a sequence which is at least 60% identical to the six nucleic acids or
CC their open reading frames, fragments of at least 15 nucleotides,
CC naturally occurring variants, or a DNA insert of the plasmid deposited
CC with the American Type Culture Collection as Accession No. not defined in
CC the specification, which encodes the amino acid sequence). Also included
CC are a host cell containing the nucleic acids (used to produce the
CC proteins), the encoded proteins, an antibody that selectively binds to
CC the polypeptide, and identifying a compound that binds to/modulates the
CC activity of the polypeptide. The nucleic acid molecules, polypeptides and
CC methods are useful for: diagnosing, treating cancer, aberrant cellular
CC proliferation and/or differentiation, immune disorders, heart disorders,
CC cardiovascular disorders including endothelial cell disorders,
CC haematopoietic disorders, blood vessel disorders, brain disorders, pain
CC and metabolic disorders, liver disorders and platelet disorders (many
CC examples of these disorders are given in the specification). The present
CC sequence is the Human cDNA 28472 encoding a glycoprotease
XX

SO Sequence 1820 BP; 543 A; 365 C; 393 G; 518 T; 0 U; 1 Other;
Alignment Scores:
Pred. No.: 8,22e-128 Length: 1820
Score: 1203.00 Matches: 232
Percent Similarity: 98.33% Conservative: 3
Best Local Similarity: 97.07% Mismatches: 4
Query Match: 97.02% Indels: 0
DB: 8 Gaps: 0
US-10-649-273-2_COPY_176_414 (1-239) x ACA60887 (1-1820)
QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlnGlyValSerLeuAspPhe 20
Db 671 CTGTTGGCATTAAGTTCAGAGAGTTTCAGATTTCGCTTCGTGAAGCTCTTGGACATA 730
QY 21 AlaProGlyAspMetLeuAspPheValAlaArgArgLeuSerLeuIleValHisProGlu 40
Db 731 GCACCAAGTGCAGCTTGACAGAGTGCAGAGACCTTCTTTAATTAACATCCAGAG 790
QY 41 CysSerThrMetSerGlyGlyValAlaIleGlnHisLeuAlaValGlnGlyValAsnArgPhe 60
Db 791 TGCCTCCACCAAGAGTGGGAGAAAGCCATAGACATTGGCCAAACAGAGAAATGATGATT 850
QY 61 HisPheAspIleValProProLeuHisAlaValAsnGlyValAspPheSerThrGly 80
Db 851 CATTTTGACATCAAACTCCCTGCATCATCTTAATAAATTTGATTTCTTTTACTGCA 910
QY 81 LeuGlnHisValIleThrAspPheIleLeuMetValGlnGlyValGlnGlyValGlnGly 100
Db 911 CTTCAACACGTTACTGATTAATAATATGAGAAACAGAGAAACAGAGAGATTTAGAGAG 970
QY 101 GlyGlnIleLeuSerSerAlaAlaAspIleAlaIleThrValGlnHisThrMetAlaCys 120
Db 971 GGGCAATCTGCTCTTCACGACGACACTTCTGCGACAGTACAGACACATGCGACT 1030
QY 121 HisLeuValValArgThrHisArgAlaIleLeuPheCysGlyGlnArgAspLeuPro 140
Db 1031 CATCTTGGAAGAAGACATCGGGCTATTCGTTTGTAAAGCAGAGACTTGTTACT 1090
QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaIleSerAsnPheTyrIleArgArg 160
Db 1091 CAAATATATGCACTGCTGTCATCTGCTGTCGAGTGCAGATCTTCATATCCGACGA 1150
QY 161 AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
Db 1151 GCTCTGAAATTTTAACCAAGCAGACACATGCTTGTGTGTGCTCTCCAGACTA 1210
QY 181 CysThrAspAsnGlyIleMetIleAlaThrAsnGlyIleGluArgLeuArgAlaGlyLeu 200
Db 1211 TGCACCTGATTAATGCGATTATATGATTCATGCAATGCTATTTGAAGACTAGCTGGCTTG 1270
QY 201 GlyIleLeuHisAspIleGlnGlyIleArgTyrGlnProLeuCysProLeuGlyValAsp 220
Db 1271 GGCATTATTAACATGACATAGAGGATCCGCTATGAAACCAAAATGTCCTTGGATAGAC 1330
QY 221 IleSerTyrGluValGlyGlnAlaSerIleValProGlnLeuLeuMetGlnIle 239
Db 1331 ATATCAAAAGAGTGGAGAGCTTCATTAAGTACACACATTAAAAATGAGATA 1387
RESULT 8
ABSS7020
ID ABSS7020 standard; cDNA, 1821 BP.
XX
AC ABSS7020;
XX
DT 30-JAN-2003 (first entry)
XX
DE cDNA encoding novel human glycoprotease 28472.
XX
KW Cancer; aberrant cell proliferation; aberrant cell differentiation;
KW breast cancer; ovarian cancer; prostate cancer; colon cancer;
KW lung cancer; immune disorder; heart disorder; cardiovascular disorder;
KW

OS	Homo sapiens.
KX	
FH	Key
FT	Location/Qualifiers
CDS	147..1391
FT	/tag = a
FT	/product= "Glycoprotease 28472"
FT	/note= "Specifically claimed in claim 1"
XN	
PN	WO2002727233-A2.
PD	
PD	03-OCT-2002.
XX	
PF	08-NOV-2001; 2001WO-US046724.
XX	
PR	08-NOV-2000; 2000US-0246768P.
PR	08-NOV-2000; 2000US-0246772P.
PR	15-NOV-2000; 2000US-0249185P.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
PI	Leihey KR, Kapeller-Libermann R, Glucksmann M;
PI	WPI; 2003-029938/02.
DR	P-PSDB; ABG71162.
XX	
PT	New adenosine deaminase, glycoprotease and seven transmembrane domain
PT	nucleic acids and polypeptides, designated 38650, 28472, 5495, 65507,
PT	81588 and 14354, useful for treating e.g. leukemias, Hodgkin's disease or
PT	hypertension.
XX	
PS	Claim 2; Fig 8A-B; 178pp: English.
CC	The invention describes isolated 38650, 28472, 5495, 65507, 81588 and
CC	14354 nucleic acid molecules (I) and their encoded polypeptides (II). The
CC	38650 nucleic acid molecule comprises a sequence encoding adenosine
CC	deaminase. The 28472 nucleic acid molecule comprises a sequence encoding
CC	a human glycoprotease. The 5495, 65507, 81588 and 14354 comprise
CC	sequences that encode a human seven transmembrane domain (7TM). The
CC	38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid and polypeptide
CC	sequences are useful for diagnosing, preventing or treating a subject
CC	with or at risk of developing a disorder, e.g. cancer or aberrant
CC	cellular proliferation and/or differentiation (e.g. breast, ovarian,
CC	prostate, colon or lung cancer), immune disorders, heart disorders,
CC	cardiovascular disorders, endothelial disorders, hematopoietic disorders,
CC	blood vessel disorders, brain disorders, pain and metabolic disorders,
CC	liver disorders or platelet disorders. These disorders include carcinoma,
CC	sarcoma, leukaemia, Hodgkin's disease, autoimmune disorders,
CC	hypertransfusion, arteriosclerosis, heart failure, myocardial infarction,
CC	ischaemic heart disease, Crohn's disease, Grave's disease, Kawasaki,
CC	syndrome, Raynaud's disease, aneurysm, cerebral ischaemia, peripheral
CC	neuropathy, Alzheimer's disease, Parkinson's disease, anorexia nervosa,
CC	cachexia or diabetes. This sequence encodes the novel human glycoproteases
CC	28472
XX	
SO	Sequence 1821 BP; 543 A; 365 C; 394 G; 518 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	8-22E-128	Length:	1821
Score:	1203.00	Matches:	232
Percent Similarity:	98.33%	Conservative:	3
Best Local Similarity:	97.02%	Mismatches:	4
Query Match:	97.02%	Indels:	0
GB:	10	Gaps:	0

[illegible]

24-JAN-2002; 2002US-0350435P.
 XX (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Iseogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuno Y;
 DR WPI: 2003-395539/38.
 P-PSDB; ADA54471.
 XX
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 XX
 PS Claim 1, SEQ ID NO 400; 205pp; English.
 XX
 CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.
 XX
 SQ Sequence 2208 BP; 660 A; 454 C; 451 G; 643 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 3,88e-111 Length: 2208
 Score: 1059.00 Matches: 211
 Percent Similarity: 88.28% Conservative: 0
 Best Local Similarity: 88.28% Mismatches: 4
 Query Match: 85.40% Indels: 24
 DB: 10 Gaps: 1
 US-10-649-273-2_COPY_176_414 (1-239) x ADA52832 (1-2208)
 QY 1 LeuLeuAlaLeuValGlnGlyValIserAspPheLeuLeuGlnGlySerLeuAspIle 20
 DB 869 CTTGTGGCATTACTGATCAAGAGTTTCAGATTTCCTCTTGAAAGCTTTTGACATA 928
 QY 21 AlaProGlyAspMetLeuAspLeuValAlaArgIleuSerLeuIleYsHsAProGlu 40
 DB 929 GCACCAAGGTGACATGCTTGACAAAGTGGCAAGAACTTCCTTAATAAATCCAGAG 988
 QY 41 CysSerThreSerGlyGlyValAlaIleGlnHisLeuAlaValGlnGlyAsnArgPhe 60
 DB 989 TGCCTCCACCATGAGTGGTGGAAAGCCATAGACATTTGGCAAAACAAAGAAATGATTT 1048
 QY 61 HisPheAspIleLeuSerProLeuHisAlaValAsnGlyAspPheSerPheThrGly 80
 DB 1049 CATTTGACATCAAACTCCCTTGCAATCATGCTMAAAATGTGATTTCTTTACTGGA 1108
 QY 81 LeuGlnHisValThrAspLysIleIleMetLysGlnGluGluGlyIleGlnLys 100
 DB 1109 CTTCAACAGCTTACTGATTAATAATTAATAAAGAAAGAAAGCAAGATTGAGAG 1168
 QY 101 GlnGlnIleuSerSerAlaAlaAspIleAlaIleThrValGlnHisThrMetAlaCys 120
 DB 1169 GGGCAATCTGTCTTGACAGACGACATTTGCTGCCACAGTACGACACAAATGGCATGT 1228
 QY 121 HisLeuValIysArgThrHisArgAlaIleLeuPheCysLysGlnAlaGAspLeuLeuPro 140
 DB 1229 CATCTGTGAAAAGACACATCGGGCTATTCTGTTTGTAAAGCAGAGACTTCTTACT 1288
 QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyValAlaSerAsnPheYrIleArgArg 160
 DB 1289 CAATAAATATGACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1348
 QY 161 AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
 DB 1349 GCTCTGGAATTTTAAACAAACGACACAGTGCATTTGTTGTCTCTCTCCACGACTA 1408
 QY 181 CysThrAspAsnGlyIleMetIleAlaIleArgAsnGlyIleGlnArgLeuArgAlaGlyLeu 200

DB 1409 TGCACATATATGACATTATGATTC----- 1435
 QY 201 GlyIleLeuHisAspIleGlnGlyIleArgYrGluProLysCysProLeuGlyValAsp 220
 DB 1436 -----TGATGCTCTTGAGTAAAC 1456
 QY 221 IleSerLysGluValGlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 239
 DB 1457 ATATCAAAAGAGTGGAGAAAGCTTCATMAAAGTACACACATTAATAATGAGATA 1513
 RESULT 10
 ADQ24627
 ID ADQ24627 standard; DNA; 2890 BP.
 XX
 AC ADQ24627;
 XX
 DT 26-AUG-2004 (first entry)
 XX
 DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 7447.
 XX
 KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
 KW ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2004048938-A2.
 XX
 PD 10-JUN-2004.
 XX
 PF 26-NOV-2003; 2003WO-US038193.
 XX
 PR 26-NOV-2002; 2002US-0429739P.
 XX
 PA (PROT-) PROTEIN DESIGN LABS INC.
 PI Aziz N, Ginsburg WM, Zlotnick A;
 DR WPI: 2004-441206/41.
 XX
 PT Early detection of soft tissue sarcoma comprises determining expression
 PT of a gene in a first soft tissue sample and a normal soft tissue sample
 PT and comparing the gene expression, also useful in treating soft tissue
 PT sarcoma.
 XX
 PS Example 2; SEQ ID NO 7447; 210bp; English.
 XX
 CC The invention relates to a novel method for detecting soft tissue sarcoma
 CC which comprises obtaining a first soft tissue sample from an individual
 CC and a normal soft tissue sample from the same or different individual,
 CC determining the expression of a gene in both samples and comparing the
 CC expression of the gene in both soft tissue samples, where a higher level
 CC of protein expression in the first soft tissue sample indicates the
 CC presence of soft tissue sarcoma. The method of the invention has
 CC cyrostatic applications and may be useful for detecting soft tissue
 CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
 CC acid sequences may be useful in diagnostic and screening applications.
 CC The current sequence is that of a human soft tissue sarcoma-upregulated
 CC DNA of the invention. The current sequence is not shown within the
 CC specification per se but was submitted in CD format by the inventor.
 XX
 SQ Sequence 2890 BP; 869 A; 609 C; 611 G; 789 T; 0 U; 12 Other;
 XX
 Alignment Scores:
 Pred. No.: 5.67e-111 Length: 2890
 Score: 1059.00 Matches: 211
 Percent Similarity: 88.28% Conservative: 0
 Best Local Similarity: 88.28% Mismatches: 4
 Query Match: 85.40% Indels: 24
 DB: 12 Gaps: 1
 US-10-649-273-2_COPY_176_414 (1-239) x ADQ24627 (1-2890)

QY 1 LeuLeuAlaLeuValGlnGlyValIserAspPheLeuLeuGlyLysSerLeuAspIle 20
 DB 1526 CCGTGGCAATGATTGTTAGAGAGTTTCAGATTTCCTGTTGGAAAGTCTTGGACATA 1585
 QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHISProGlu 40
 DB 1586 GCAACGAGTGACATGCTTGAACAGGTGGCAAGAAAGACTTCTTTATATAACATCCAGAG 1645
 QY 41 CysSerThrMetSerGlyGlyLysAlaIleGlnHISLeuAlaLysGlnGlyAsnArgPhe 60
 DB 1646 TCGTCACCATGATGCTGTGGAAAGCATTAACATTGGCCAAAGAAAGAAATAGATT 1705
 QY 61 HisPheAspIleLysProProLeuHisHisAlaLysAsnGlyAspPheSerPheThrGly 80
 DB 1706 CATTTTGACATCAACCTCCCTGCATCATGCTAAATAATGATGATTTCTTTTACTGGA 1765
 QY 81 LeuGlnHisValThrAspLysIleIleMetLysLysGlnLysGlnGlyIleGluLys 100
 DB 1766 CTTCAACAGATTATCTATTAATAATTAATGAAGAAAGAAAGAAAGATTGAGAGAG 1825
 QY 101 GlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCys 120
 DB 1826 GGGCAAAATCTGCTTCAGACAGACATTCCTGCCACATACAGACACAAATGGCATGT 1885
 QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuPro 140
 DB 1886 CATCTTGTAAGAAAGACATCGGGCATTCCTGTTTGTAGACAGAGACTTGTATCCT 1945
 QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPheIleArgArg 160
 DB 1946 CAATAAATAAGCAGATCTGTTGATCATCTGGTGGTCCAGATTAATCTATATCCGAGA 2005
 QY 161 AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
 DB 2006 GCTCTGGAATTTTCAACAGACAGACAGCATTCCTGTTGTCCTCCACAGACTA 2065
 QY 181 CysTrpAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeu 200
 DB 2066 TGCACATGATATGAGCATTAATGATGCA----- 2092
 QY 201 GlyIleLeuHisAspIleGluGlyIleArgGlyGluProLysCysProLeuGlyValAsp 220
 DB 2093 -----TGATGTCCTCTTGGAGTGAAC 2113
 QY 221 IleserLysGluValGlyGlnAlaSerIleLysValProGlnLeuLysMetGlnIle 239
 DB 2114 ATATCAAAAGAGTTGAGAGAGCTTCATTAAGATACCAATTAATAATGAGATTA 2170
 RESULT 11
 ADE31345/c
 ID ADE31345 standard; DNA; 3358 BP.
 XX ADE31345;
 XX 29-JAN-2004 (first entry)
 DE Human diagnostic and therapeutic polynucleotide (dthp), SEQ ID NO 100.
 XX
 XX diagnostic and therapeutic polynucleotide; dthp; antiarteriosclerotic;
 XX antiinflammatory; cerebroprotective; antilipemic; antidiabetic;
 XX immunosuppressive; neuroprotective; nootropic; neuroleptic; tranquilizer;
 XX osteopathic; antirheumatic; antirheumatic; cyrostatic; hepatotropic;
 XX virucide; haemostatic; anti-HIV antithyroid; thymomimetic;
 XX dermatological; antibacterial; fungicide; antiparasitic; anticonvulsant;
 KW thrombolytic; anticoagulant; anorectic; vasotropic; antulser;
 KW gene therapy; protein replacement therapy; human; gene; ds.
 OS Homo sapiens.
 XX
 XX WO2003062376-A2.
 XX 31-JUL-2003.

PF 13-JAN-2003; 2003WO-US001096.
 XX
 PR 16-JAN-2002; 2002US-0349384P.
 PR 17-JAN-2002; 2002US-0349413P.
 PR 17-JAN-2002; 2002US-0349946P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Jones AL, Dahl CR, Gietzen D, Chinn J, Dufour GR, Jackson JL,
 PI Yu JY, Tuason O, Yap PE, Amshy SR, Dam TC, Liu TF, Gerstin EH,
 PI Petralia CH, Lewis SA, Chen A, Marwaha R, Lan KY, Urashka ME,
 PI Kristnam SR, Kolluru V, Panesar IS;
 XX
 DR WPI; 2003-636732/60.
 DR P-PSDB; ADE31156.
 XX
 PT New human diagnostic and therapeutic polynucleotides and polypeptides,
 PT useful for diagnosing, treating or preventing e.g. leukemia, brain
 PT cancer, atherosclerosis, AIDS, thyroiditis, infections, obesity, stroke
 PT or Alzheimer's.
 PS Claim 1; SEQ ID NO 100; 634pp; English.
 CC The invention relates to a novel isolated human diagnostic and
 CC therapeutic polynucleotide (designated dthp). The novel dthp
 CC polynucleotide comprises: any of 188 DNA sequences consisting of 195-7798
 CC base pairs fully defined in the specification; a polynucleotide
 CC comprising a naturally occurring polynucleotide sequence at least 90%
 CC identical to the dthp polynucleotide; a polynucleotide complementary to
 CC the dthp polynucleotide or its polynucleotide which is at least 90%
 CC identical; or an RNA equivalent of any of the polynucleotides mentioned
 CC above. The dthp polynucleotides have the following activities:
 CC antiarteriosclerotic, antiinflammatory, cerebroprotective, antilipemic,
 CC antidiabetic, immunosuppressive, neuroprotective, nootropic, neuroleptic,
 CC tranquilizer, osteopathic, antirheumatic, antirheumatic, cyrostatic,
 CC hepatotropic, virucide, haemostatic, anti-HIV, antithyroid, thymomimetic,
 CC dermatological, antibacterial, fungicide, antiparasitic, anticonvulsant,
 CC thrombolytic, anticoagulant, anorectic, vasotropic, and antulser. The
 CC novel dthp polynucleotides polypeptide can be used in gene therapy and
 CC protein replacement therapy. The dthp polynucleotides or dthp
 CC polypeptides are useful for diagnosing, preventing or treating diseases
 CC associated with the expression of human molecules. In particular, these
 CC diseases include cancers (e.g. adenocarcinoma, leukaemia, melanoma, brain
 CC cancer), breast cancer, cervix cancer, bone cancer, liver cancer, lung
 CC cancer) or other cell proliferative disorders (e.g. arteriosclerosis,
 CC atherosclerosis, cirrhosis, hepatitis, polycythemia vera, primary
 CC thrombocytopenia), autoimmune/inflammatory disorders (e.g. AIDS,
 CC Addison's disease, thyroiditis, Crohn's disease, Graves' disease,
 CC Hashimoto's thyroiditis, multiple sclerosis, osteoarthritis, rheumatoid
 CC arthritis, scleroderma, systemic lupus erythematosus), infections (e.g.
 CC viral, bacterial, fungal or parasitic infection), developmental disorders
 CC (e.g. Cushing's syndrome or epilepsy), endocrine disorders (e.g. diabetes
 CC thrombosis, hypopituitarism, hypogonadism, gigantism, goiter) metabolic
 CC disorders (e.g. hypercholesterolaemia, hypoglycaemia, diabetes,
 CC hyperlipidaemia, obesity), neurological disorders (e.g. ischaemic
 CC cerebrovascular disease, stroke, Alzheimer's disease, Pick's disease,
 CC Huntington's disease, Parkinson's disease, Creutzfeldt-Jakob disease,
 CC anxiety, schizophrenia), gastrointestinal disorders (e.g. ulcers),
 CC transport disorders (e.g. akinesia or multidrug resistance), or
 CC connective tissue disorders (e.g. Paget's disease or rickets). This
 CC polynucleotide sequence represents one of the human dthp DNA sequences
 CC of the invention.
 XX
 SQ Sequence 3358 BP; 1105 A; 577 C; 601 G; 1075 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,4e-103 Length: 3358
 Score: 995.50 Matches: 209
 Percent Similarity: 87.82% Conservative: 0
 Best Local Similarity: 87.82% Mismatches: 2
 Query Match: 80.28% Indels: 29
 DB: 10 Gaps: 1

US-10-649-273-2_COPY_176_414 (1-239) x ADR31345 (1-3358)			
QY	29	ValAlaArgAGLeuSerLeuIleLysHisProGluCysSerThrMetSerGly	48
Db	2104	GTGGCAGAAACATTTCTTTATATAAACATCCAGAGTCTCCACCATAGTGGGGAA	2045
QY	48	SAAlIleGluHisLeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLe	68
Db	2044	AGCCATCAACACATTTGGCCAAACAGAAATAGATTCAATTTTGACATCAAACTCCCTT	1985
QY	68	uHisHisAlaLysAsnCyAspPheSerPheThrGlyLeuGlnHisValThrAspLysI	88
Db	1984	GCATCATGCTAAAAATTGTGATTTTCTTTTAAGTGAAGTCTCAACAGCTTACTGATTAAT	1922
QY	88	eIleMetLysLysGluLysGluGlu-----	96
Db	1924	AATATATGAAAAAGAAAAAGAAAGTATATTCTTAATTAGTAAAGTTGACAGATAA	1865
QY	97	-----GlyIleGluLysGlnI	103
Db	1864	TATTCCTGGATTGGCTTAAAAATAGCTGCTCATTTCTGCAGAGTATTGAAAGGGGCAAA	1800
QY	103	IeLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCysHisLeuV	123
Db	1804	TCCTGTCTTCAGCAGCAGACACATGCTGGCAGCAGTACAGACACATAGGATGTCATCTTG	1745
QY	123	AlLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuProGlnAsnA	143
Db	1744	TGAAAGAAGACATCGGGCTATTCTGTTTGTAAACAGAGAGACTGTGTAACCTCAAAATA	1685
QY	143	SPAlaValIleuValAlaSerGlyGlyValAlaSerAsnPheThrIleArgArgAlaLeuG	163
Db	1684	ATGCAGTACTGGTTCATCTGT-GGTGTCCGCAAGTAACCTCTATATCCGACAGCTCTGG	1622
QY	163	IuIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeuCysThrA	183
Db	1625	AAATTTTAAAC-AAGCACAACAGTGCACATTTGTGTGTCTCTCCACATATGACACTG	156
QY	183	SPAsnGlyIleMetIleAlaThrPaanglyIleGluArgLeuArgAlaGlyLeuGlyIleL	203
Db	1566	ATAATGGCATTATGATGTGATGAGATGGATTAAGAAAGACTAGTCTGGCTTGGGCAATT	150
QY	203	euhIleAspIleGluGlyIleArgGlyTyrGluProLysCysProLeuGlyValAspIleSerL	223
Db	1506	TACATGACATAGAAAGCATTCCCTATGAACCAAAATGTCCTTGTGAGTACATATACA	144
QY	223	YSGluValGlyGluAlaSerIleLysValProGlnLeuLysMetGluIle	239
Db	1446	AAAGAATTGGAGAACCTTCATTAAGTACCAACATTAAAAATGGAGATA	1397
RESULT 12			
ABO75508			
ID ABO75508 standard; DNA; 1572 BP.			
XX	AC	ABO75508;	
XX	XX		
XX	DT	07-NOV-2002 (first entry)	
XX	XX		
DE	XX	Murine sialoglycoprotease-like gene sequence SEQ ID NO:7.	
XX	XX		
KW	XX	Murine; mouse; protease; calcium activated neutral protease type 5;	
KW	XX	CAMP5; cryptase 4; sialoglycoprotease; enzyme; genetic disease;	
KW	XX	neurological; neuropsychological; psychotic illness; transgenic animal;	
XX	XX	gene; ds.	
OS	XX	Mus musculus.	
XX	XX		
XX	XX	WO200245491-A2.	
XX	XX		
XX	XX	13-JUN-2002.	
XX	XX		
XX	XX	05-DEC-2001; 2001WO-US046405.	
XX	XX		

XX	PA	(DELT-) DELTAGEN INC.
XX	PB	
XX	PC	Allen KD, Lewiten MW;
XX	PD	WPI, 2002-657389/70.
XX	PE	
XX	PF	Novel transgenic animal, comprising a disruption in protease target gene,
XX	PG	is useful for identifying agents that ameliorates a phenotype associated
XX	PH	with a disruption in a protease target gene.
XX	PI	
XX	PJ	Example 3; Fig 7; 62pp; English.
XX	PK	
XX	PL	The present invention describes a non-human transgenic animal (I)
XX	PM	comprising a disruption in a protease target gene (Pg) selected from
XX	PN	calcium activated neutral protease type 5 (CAPN5) gene, trypsin 4 gene
XX	PO	and siatologlycoprotease-like gene. Also described is a targeting construct
XX	PP	(II), comprising a first polynucleotide sequence homologous to at least a
XX	PQ	first portion of Pg, a second polynucleotide sequence homologous to at
XX	PR	least a second portion of Pg and a selectable marker. (II) is useful for
XX	PS	producing a transgenic mouse comprising a disruption in a protease target
XX	PT	gene, by introducing (II) into a cell, introducing the cell into a
XX	PV	blastocyst, implanting the resulting blastocyst into a pseudopregnant
XX	PW	mouse, where the pseudopregnant mouse gives birth to a chimeric mouse,
XX	PX	and breeding the chimeric mouse to produce the transgenic mouse. (I) is
XX	PY	useful for identifying an agent that modulates the expression or function
XX	PZ	of a protease target gene, by administering an agent to (I) and
XX	QA	determining whether the expression or function of the disrupted protease
XX	QB	target gene in (I) is modulated. (I) is also useful for testing the
XX	QC	efficacy of proposed genetic and pharmacological therapies for human
XX	QD	genetic diseases, such as neurological, neuropsychological or psychotic
XX	QE	illnesses. The present sequence represents murine siatologlycoprotease-like
XX	QF	gene sequence, which is used in an example from the present invention
XX	QG	
XX	QH	SQ Sequence 1572 BP; 455 A; 337 C; 340 G; 429 T; 0 U; 7 Other:
XX	QI	
XX	QJ	Alignment Scores:
XX	QK	Pred. No.: 1.28e-89 Length: 1572
XX	QL	Score: 870.00 Matches: 184
XX	QM	Percent Similarity: 82.77% Conservative: 13
XX	QN	Best Local Similarity: 77.31% Mismatches: 37
XX	QO	Query Match: 70.16% Indels: 5
XX	QP	DB: 6 Gaps: 2
XX	QQ	
XX	QR	US-10-649-273-2_COPY_176_414 (1-239) X ABQ75508 (1-1572)
XX	QS	
XX	QT	QY 1 LeuleuAlaleuValGInglYValSerAapPheLeuLeuLGIyLySerLeuAaePile 20
XX	QU	DB 528 CTGTGGAKRTTGAAGTCCTCAAGSTTTCCGATTTCCGTCCTCTGGGAAGCTTTTGACATA 567
XX	QV	21 AAAPrGLYAspMetLeuAspLySValAlAAcGaRgLeuSerLeuLleLyShISProGlu 40
XX	QW	DB 588 GCACCAAGGCGACAATGCTTTCACAGAGTGGCAGAAGACTTTCTTAATCAAACATCCAGAA 647
XX	QX	41 CySSerThMeTSerClYGlYLySaLaIleGIuHISLeuAlalYSGLInglYaanaRqPe 60
XX	QY	DB 648 TCTTCAACAATAGTGGTGGAAAAGCTATAGAACAATTGCCCAAGAAGCAATAATGATTTC 707
XX	QZ	61 HisPheAepILlalySProFroLeuHISHISAlalvSaSnCYeaSPheSerPHeTrnGly 80
XX	RA	DB 708 CATTTTACATCAATCCACCTATAGCAGAAATGCCTMAAATYGCAATTTTCTTACACGGGA 767
XX	RB	81 LeuGINHISValThrASpLySLlelleMetyLySGlyLysGLInglUGlyIleGIuLyS 100
XX	RC	DB 768 CTTCAACATATTAAGTAACTAAATAACAACACAAAGAAAAMAAGAAGCATTTGAGAG 827
XX	RD	101 GIyGINIlleLeuSeSerAlaaIAaePILleaIaIeThrVaIGlnHISThrMetAlaCYs 120
XX	RE	DB 828 GGAGCAATCTGTCACTCACTGCAGACATTGCTGCTGCGGTACAGATGCAAGGCTGC 887

QY	121	HisLeuValIysAspGlyThrIleArgAlaIleLeuPheCysGlnArgAspLeuPro	140
Db	888	CACCTTGGGAAAAGACACATCGCCCTATTCTGTTTCCACGACGAAAAATTGCTCT	947
QY	141	GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPheTyrIleArgArg	160
Db	948	CCACCTAAACGACGATATTGTTGTTACTCGAAGGTGGTTCGAAGTAACCTGTACATCCGAAAA	1007
QY	161	AlaIleuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLeu	180
Db	1008	GCATTGGAAATGTCGCAAAATGCACGCGATGCACCGTGTGTGTGCCACT-TCAGACTG	1066
QY	181	CysThrAspAsnGlyIleMetIleAlaTyrAsnGlyIleGluArgLeuArgAlaGlyLeu	200
Db	1067	TGCATCGACATATGGCGCATATGATTTCATGAAATGCAGATTCAGTCCCTGCTT	1126
QY	201	GlyIleLeu-HisAsp-IleGluGlyIleArgTyrGluProLysCysProLeuGlyVal-	219
Db	1127	GGCCTTTTACCATATGATGATAGACATTCGGTTATTACCCAAATGTCCTCTTTGAGTGA	1186
QY	220	--AspIleSerLysGluValGlyGlyAlaSerIleLysValProGlnLeu	235
Db	1187	GGCATTTATCCGAAAGAAAGTTGGCAGAA--AGCTTGCCTATTAAAAAGTTA	1233
RESULT 13			
AA	AS84622	standard; cDNA, 2734 BP.	
ID	AS84622;		
AC	AS84622;		
XX			
DT	13-FEB-2002	(first entry)	
XX			
DE	DNA encoding novel human diagnostic protein #20426.		
XX			
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	MO200175067-A2.		
PD			
XX			
PF	11-OCT-2001.		
XX			
PR	30-MAR-2001; 2001WO-US008631.		
XX			
PR	31-MAR-2000; 2000US-00540217.		
PR	23-AUG-2000; 2000US-00649167.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
P1	Drmnac RT, Liu C, Tang YT;		
XX			
DR	WPI; 2001-639362/73.		
XX			
DR	P-PSDB; ABG20435.		
XX			
PT	New isolated polynucleotide and encoded polypeptides, useful in		
PT	diagnostics, forensics, gene mapping, identification of mutations		
PT	responsible for genetic disorders or other traits and to assess		
PT	biodiversity.		
XX			
PS	Claim 1; SEQ ID NO 20426; 103pp; English.		
XX			
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)		
CC	sequences. (I) is useful as hybridisation probes, polymerase chain		
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,		
CC	and in recombinant production of (II). The polynucleotides are also used		
CC	in diagnostics as expressed sequence tags for identifying expressed		
CC	genes. (I) is useful in gene therapy techniques to restore normal		
CC	activity of (II) or to treat disease states involving (II). (II) is		
CC	useful for generating antibodies against it, detecting or quantitating a		
CC	polypeptide in tissue, as molecular weight markers and as a food		
CC	supplement. (II) and its binding partners are useful in medical imaging		
CC	of sites expressing (II). (I) and (II) are useful for treating disorders		

CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC	coding sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 2734 BP; 773 A; 545 C; 639 G; 763 T; 0 U; 14 Other;
XX	Alignment Scores:
XX	Pred. No.: 4.02e-58 Length: 2734
XX	Score: 599.00 Matches: 131
XX	Percent Similarity: 71.57% Conservative: 10
XX	Best Local Similarity: 66.50% Mismatches: 22
XX	Query Match: 48.31% Indels: 34
XX	DB: 5 Gaps: 6
XX	US-10-649-273-2_COPY_176_414 (1-239) x AAS84622 (1-2734)
OY	43 ThrmetSerGlyGlyLysAlaIleGluHisLeuAlaLysGlnGlyAsnArgPheHisPhe 62
DB	206 ACCCTGCTAGGGGGGGCGCT-----AGATTCCAT--- 235
OY	63 AspileLysProProLeuHisHisAlaLys-----AsnCys----- 74
DB	236 -----CCTAATTTCTCCAGTGAAGATGACAGGCTGTGTAGGAAGCTGGCAGCA 283
OY	75 -----AsPheSerPheThrGlyLeuGlnHisValIThrAsp 86
DB	284 TCTGCTTGGGAGAGCCACGAGGACTTTTACTCATGGCGGAAGGCAAAATGGGAGCTAG 343
OY	87 LysIleIleMetLysLysGluLysGlu-----GlyIleGluLysGlyGlnIleLeu 104
DB	344 CGCTTCAATGCGCAGGAGCAGGACCGAGAGAGGGGGGATTTGAGAAAGGGCAAACTCTG 403
OY	105 SerSerAlaAlaAspIleAlaIleAlaThrValGlnHisThrMetAlaCysHisLeuValLys 124
DB	404 TCTTACAGCAGCAGACATTGTCTGCCACAGTACAGCACACAAATGGCATTGATCTTGAGAA 463
OY	125 ArgThrHisIleAlaIleLeuPheCysLysGlnIleGAspLeuLeuProGlnIleAsnAla 144
DB	464 AGAACACATCGGGCTAATCTGTTTGTGAAGCAGAGACTGTGTACTCAAAATTAATGCA 523
OY	145 ValIleuValAlaSerGlyGlyValAlaSerAsnPheTyrlleArgTrpAlaLeuGluIle 164
DB	524 GTACTGGTGCATCTGGTGGTGTCCCAAGTAACTTTATATACCGCAGAGCTCTGGAAATT 583
OY	165 LeuThrAsnAlaThrGlnCysThrLeuLeuCysProProProlArgLeuCysThrAspAsn 184
DB	584 TTACAAACGACACACAGTGCCTTTGTTGTCTCCCTCCACAGATTAGCACTGATTAAT 643
OY	185 GlyIleMetIleAlaIleAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHis 204
DB	644 GGCAATTAGAATGATGATGAATGGTATTTGAAGAACACTACGTGGCTGGCTTGGGCATTTAAT 703
OY	205 AspIleGluGlyIleArgTyrgluProLys-----CysProLeuGlyVal 219
DB	704 GACATGAGAGGAGCATCCGCTATGACCAAAATGATGTGGCTCTTCAGAGGCGTG 754
XX	RESULT 14
XX	ADL86725
XX	ADL86725 standard; DNA; 371 BP.
XX	ADL86725;
XX	20-MAY-2004 (first entry)
XX	DNA up-regulated in murine multipotent progenitor cells Segid 3118.

XX	gene potential; multi-lineage; cell commitment; haematopoietic stem cell;
KM	HSC; multipotent progenitor; MPP; common lymphoid progenitor; CLP;
KW	common myeloid progenitor; CMP; bone marrow stem cell; mouse; murine; ds.
XX	
OS	Mus sp.
XX	
PN	WO2003093445-A2.
XX	
PD	13-NOV-2003.
XX	
PF	05-MAY-2003; 2003WO-US014114.
XX	
PR	03-MAY-2002; 2002US-0377383P.
XX	
XX	
PA	(STOW-) STOWERS INST MEDICAL RES.
XX	
PI	L1 L;
XX	
DR	WPI; 2004-022656/02.
XX	
PT	Classifying an unknown multi-lineage affiliated gene comprises isolating
PT	expressed nucleic acid sequences from the discrete cell sub-populations.
XX	
PS	Claim 8; SEQ ID NO 3118; 123pp; English.

This invention relates to a novel method for predicting gene potential by associating nucleic acid sequences of unknown function with particular sub-population profiles. Specifically, it refers to classifying an unknown multi-lineage affiliated gene by collecting hybridisation data to develop a gene expression map, in order to determine the discrete sub-population where it is expressed. The present invention describes methods for predicting the lineage commitment of genes associated with the self-renewing haematopoietic (blood) stem cells (HSCs), as well as the non-self-renewing multipotent progenitors (MPPs), common lymphoid progenitors (CLPs) and common myeloid progenitors (CMPs), which are collectively referred to as bone marrow stem cells populations. As such, these methods can be used to identify associated multi-lineage affiliated genes and hence the underlying molecular mechanisms in physiological haematopoietic development. This polynucleotide sequence is DNA associated with a murine MPP sub-population of cells of the invention.

SQ Sequence 371 BP; 104 A; 75 C; 84 G; 92 T; 0 U; 16 Other;

Alignment Scores:	
Pred. No.:	2,75e-44
Score:	468.00
Percent Similarity:	78.86%
Best Local Similarity:	74.80%
Query Match:	37.74%
DB:	12
Length:	377
Matches:	92
Conservative:	5
Mismatches:	26
Indels:	0
Gaps:	0

US-10-649-273-2_COPY_176_414 (1-239) x ADL86725 (1-371)

Qy	109	AspIleAlaIaAlaThrValGlnIsthrMetAlaCysHisIleuValIysAspThrHisLeuArg	128
Db	2	GACATTCGCTGCTGGGTACAGCAGTACGACAGCTGTCACCTTCGCAAAAGAACACATCCG	61
Qy	129	AlaIleIleuPheCysIysGlnArgAspLeuIleuProGlnAsnAsnAlaValIleuValAla	148
Db	62	GCTATTCCTGTTTGGCAGCAGAAATTTGCTCTCTCCAGCTAAACGACAGTATTAGTTGTA	121
Qy	149	SerGlyGlyValAlaSerAsnPheTyrIleArgArgAlaIeuGlyIleIeuThrAsnAla	168
Db	122	TCTGGAGGTGTTGCAGATGATCTTGTAACATCCGAAAGACATTCGAAATTTGTCGCAAAATGCA	181
Qy	169	ThrGlnCysThrIleuIeuCysProProArgIeuCysThrAspAsnGlyIleMetIle	188
Db	182	ACGCAGTGCACCTTTGTTGTGCCACTCCAGACTGTGCATGACATATGATCAT	241
Qy	189	AlaIlePheAsnGlyIleGlnArgIeuArgAlaGlyIleuGlnIleIeuHisAspIleGly	208
Db	242	GCATGAAATGGAAATTCAGATTCACGTCGACGCTTGGCACTTTATCATCTGAATTAAGAAAC	301

QY	209	IIeaGTyGtUpProLyeCyAProLeugIyVAlAspIleSeTySGtUyAla	225
Db	302	ATCCNNATGNACCAAAATNTCNTNTTGTGAGTAGACATNTCCAGAGAGTTCGAGAAGCT	361
QY	229	SeTtleTyS 231	
Db	362	GCCTATPAA 370	

RESULT 15
ADL86726
ID ADL86726 standard; DNA; 371 BP.

DNA up-regulated in murine multipotent progenitor cells SeqID 3119.

KW gene potential; multi-lineage; cell commitment; haematopoietic stem cell;
KW HSC, multipotent progenitor; MPP: common lymphoid progenitor; CLP:
KW common myeloid progenitor; CMP: bone marrow stem cell; mouse; murine; ds

OS	Mus sp.
XX	
PN	W02003093445-A2.

PF 05-MAY-2003; 2003WO-US014114.
XX
PR 03-MAY-2002; 2002US-0377383P.
XX
PA (STOW-) STOWERS INST MEDICAL RES

This invention relates to a novel method for predicting gene potential by associating nucleic acid sequences of unknown function with particular sub-population profiles. Specifically, it refers to classifying an unknown multi-lineage affiliated gene by collecting hybridisation data to develop a gene expression map, in order to determine the discrete sub-population where it is expressed. The present invention describes methods for predicting the lineage commitment of genes associated with the self-renewing haematopoietic (blood) stem cells (HSCs), as well as the non-self renewing multipotent progenitors (MPPs), common lymphoid progenitors (CLPs) and common myeloid progenitors (CMPs), which are collectively referred to as bone marrow stem cell populations. As such, these methods can be used to identify associated multi-lineage affiliated genes and hence the underlying molecular mechanisms in physiological haematopoietic development. This polynucleotide sequence is DNA associated with a murine MPP sub population of cells of the invention.

SQ Sequence 371 BP; 104 A; 75 C; 84 G; 92 T; 0 U; 16 Other;

Designated record:	
Pred. No.:	2.75e-44
Score:	468.00
Percent:	78.86%
Best Similarity:	74.80%
Best Local Similarity:	37.74%
Query Match:	12
DB:	0
Length:	371
Matches:	92
Conservative:	5
Mismatches:	26
Indels:	0
Gaps:	0

US-10-649-273-2 COPY 176 414 (1-239) X ADL86726 (1-371)

109 AspllealalarnvalGlnHsrHmetalaCysHslsleuVallysarGThrhIarg 128

Db 2 GACATTGCTGCGTGGCGGTACAGCATGCAACAGGCTGCACCTTGCGAAAGAAACATCCG 61
 Qy 129 AlaIleLeuPheCysElyGlnArgAspLeuProGlnAsnAsnAlaValLeuValAla 148
 Db 62 GCTATTCTGGTTTGGACGAGAGAAATTGCTCTCCAGCTAAACCGATATTAAGTGTGA 121
 Qy 149 SerGlyGlyValAlaSerAsnPheYrIleArgArgAlaLeuGluIleLeuThrAsnAla 168
 Db 122 TCTGGAGGGTGGCAAGTAACTGTGACATCCGAAAGCACTGGAAATTTGCGCAATGCA 181
 Qy 169 ThrGlnCysThrLeuLeuCysProProArgLeuCysThrAspAsnGlyIleMetIle 188
 Db 182 ACGCAGTGCACGTTGTGTGTCCACTCCAAAGCTGTGACACTGACATGGCATCATGATT 241
 Qy 189 AlaIArgAsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGly 208
 Db 242 GCATGGAATGGAAATTGAANGATTACGTGCGNGCTTGGCCTNTTACNTGANNATGAAGAC 301
 Qy 209 IleArgIYrGlnProIYcCysProLeuGlyValAspIleSerIYcGluValGlyIuAla 228
 Db 302 ATCCNNNTTGNACCAAAATNTCNTNTYTGGAGTAGACATNTCCAGAGAAGCTTGAGAAGCT 361
 Qy 229 SerIleIYs 231
 Db 362 GCCATPAAA 370

Search completed: June 16, 2005, 20:15:17
Job time : 354.511 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 16, 2005, 18:30:49 ; Search time 118.721 Seconds
(without alignments)
3294.036 Million cell updates/sec

Title: US-10-649-273-2_COPY_176_414

Perfect score: 1240
Sequence: 1 L1ALVQGVSDFFLLGKSLDI.....DISKEVGEASIKVPOLKMEI 239

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 segs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlp
-Q=/cg2n_1/USPOL.spool.p/US10649273/runat_15062005_111417_6057/app.query.fasta_1.1429
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.ccd1
-LIST=45 -DOCALIGN=200 -NOR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=p2n -NOR SCORE=HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10649273@cgc1_1_177 @runat_15062005_111417_6057 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA.*
1: /cg2n_6/prodata/1/ina/5A.COMB.seq.*
2: /cg2n_6/prodata/1/ina/5B.COMB.seq.*
3: /cg2n_6/prodata/1/ina/6A.COMB.seq.*
4: /cg2n_6/prodata/1/ina/6B.COMB.seq.*
5: /cg2n_6/prodata/1/ina/PTUS.COMB.seq.*
6: /cg2n_6/prodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1240	100.0	1416	4	US-09-774-528-177 Sequence 177, App
2	1240	100.0	1526	4	US-10-067-443-23 Sequence 23, Appl
3	1240	100.0	2197	4	US-10-067-443-1 Sequence 1, Appl
4	1213.5	97.9	1387	4	US-10-067-443-21 Sequence 21, Appl
5	1213.5	97.9	14364	4	US-10-067-443-20 Sequence 20, Appl
6	248.5	20.0	94750	4	US-09-596-002-38 Sequence 38, Appl
7	247	19.9	1053	4	US-09-540-236-806 Sequence 806, App
8	231	18.6	1059	4	US-09-252-991A-884 Sequence 884, App
9	231	18.6	1206	4	US-09-252-991A-801 Sequence 801, App
10	224	18.1	1830121	4	US-09-557-884-1 Sequence 1, Appl
11	224	18.1	1830121	4	US-09-643-990A-1 Sequence 1, Appl
12	221.5	17.9	996	4	US-09-902-540-6612 Sequence 6612, Ap

C	13	221.5	17.9	2582	4	US-09-902-540-503	Sequence 503, App
	14	212	17.1	1074	4	US-09-543-681A-2341	Sequence 2341, Ap
	15	205	16.5	1032	4	US-09-489-039A-2050	Sequence 2050, Ap
	16	200	16.1	1315	1	US-08-087-797-1	Sequence 1, Appl
	17	197	15.9	1092	4	US-09-107-532A-2955	Sequence 2955, Ap
	18	190	15.3	1008	3	US-08-987-121A-5	Sequence 5, Appl
	19	190	15.3	1011	3	US-08-987-121A-3	Sequence 3, Appl
	20	189	15.2	822	4	US-09-710-279-127	Sequence 727, App
	21	189	15.2	3993	4	US-09-710-279-3985	Sequence 3985, Ap
	22	187	15.1	1155	4	US-09-602-777A-147	Sequence 147, App
	23	186.5	15.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl
	24	186.5	15.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl
	25	186	15.0	1006	3	US-08-961-083-51	Sequence 51, Appl
	26	186	15.0	1006	4	US-09-536-784-51	Sequence 51, Appl
	27	186	15.0	1107	3	US-09-134-001C-1072	Sequence 1072, Ap
	28	186	15.0	10974	3	US-08-961-527-214	Sequence 214, App
C	29	184	14.8	1011	3	US-09-066-512-1	Sequence 1, Appl
	30	182	14.7	1011	4	US-09-583-110-2196	Sequence 2196, Ap
	31	182	14.7	1663	4	US-09-620-312D-6	Sequence 6, Appl
	32	181.5	14.6	432	4	US-09-328-352-261	Sequence 261, App
	33	180.5	14.6	1664976	4	US-08-916-421B-1	Sequence 1, Appl
	34	180.5	14.6	1664976	4	US-09-692-570-1	Sequence 1, Appl
	35	178	14.4	640681	4	US-09-790-988-1	Sequence 1, Appl
	36	177	14.3	1011	4	US-09-107-433-1618	Sequence 1618, Ap
	37	174.5	14.1	1230230	4	US-09-198-452A-1	Sequence 1, Appl
	38	174.5	14.1	1230230	4	US-09-438-185A-1	Sequence 1, Appl
	39	174	14.0	1026	3	US-09-149-624-1	Sequence 1551, Ap
	40	173	14.0	1101	4	US-08-956-171E-102	Sequence 102, App
C	41	170	13.7	15249	4	US-08-781-966A-102	Sequence 102, App
C	42	170	13.7	15249	4	US-08-781-966A-102	Sequence 102, App
C	43	168	13.5	3064	3	US-09-221-017B-194	Sequence 794, App
C	44	159.5	12.9	42325	4	US-08-311-731A-131	Sequence 131, App
	45	145.5	11.7	3215	4	US-09-710-279-3566	Sequence 3566, Ap

ALIGNMENTS

RESULT 1

US-09-774-528-177
Sequence 177, Application US/09774528

Patent No. 6743619
GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Yang, Yonghong
APPLICANT: Xue, Aidong J.
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Wang, Dunrui
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6743619el Nucleic Acids and
FILE REFERENCE: 802 Polypeptides
CURRENT FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 441
SOFTWARE: pt_fl_genes Version 2.0
SEQ ID NO 177
LENGTH: 1416
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (205) .. (1305)
US-09-774-528-177

Alignment Scores:

Pred. No.:	1.02e-157	1416
Score:	1240.00	239
Percent Similarity:	100.00%	0
Best Local Similarity:	100.00%	0
Query Match:	100.00%	0
DB:	4	0
	Gaps:	0

US-10-649-2_COPY_176_414 (1-239) x US-09-774-528-177 (1-1416)

QY	1	LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle	20
Db	586	CTGTTGGCACTTAAGTTCAGAGAGTTTCAAGATTTTCTGCTTTGGAAAGCTTTGGACATTA	645
QY	21	AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu	40
Db	646	GCACCAAGTGAACATGCTTGAACAAGTGGCAAGAACCTTTCTTTATATAAACATCCAGAG	705
QY	41	CysSerThrMetSerGlyGlyLysAlaIleGluHisLeuAlaLysGlnGlyAsnArgPhe	60
Db	706	TGCTCCACCATGATGATGGTGGAAAGCCATAGAACATTTGGCCAAACAAAGAAATATGATTT	765
QY	61	HisPheAspIleLysProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGly	80
Db	766	CATTTTGACATCAAACTCCCTTGTCATCATCTATAAAATGATGATTTTCTTTTACTGGA	825
QY	81	LeuGlnHisValThrAspLysIleIleMetCysLysGluLysGluGluGlyIleGluLys	100
Db	826	CTTCAACACAGTTTACTGATATAATATATATGAAAAAGAAAAAGAGAGATATTGAGAG	885
QY	101	GlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCys	120
Db	886	GGGCAAAATCTCTGCTTCCAGCAGCAGACATTCCTGTCACAGTACAGACACAAATGCGATGT	945
QY	121	HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro	140
Db	946	CATCTTGTGAAAAAGAACACATCGGGCTAATCTGTTTGTAAACAGACAGACATTTGTAACCT	1005
QY	141	GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPheTyrIleArgArg	160
Db	1006	CAAAATTAAGCAGTACTGCTGTGCACTGCTGCTGTGCGAAGTAACTTCTATATCCGACGA	1065
QY	161	AlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu	180
Db	1066	GCTCTGGAATTTTAAACAACGACACACAGTGCATTTGTTGTGTCTCTCCCAACACTA	1125
QY	181	CysThrAspAsnGlyIleMetIleAlaTyrAsnGlyIleGluArgLeuArgAlaGlyLeu	200
Db	1126	TGCACATGATTAATGGCATTTATGATTCATGATGATGATTAAGAAAGACTAGTGTGCTGCT	1185
QY	201	GlyIleLeuHisAspIleGluGlyIleArgTyrGluProLysCysProLeuGlyValAsp	220
Db	1186	GGCATTTTAACTATACATAGAGAGCATCGCTATGAAACAAATATGCTCTTTGGAGTAGAC	1245
QY	221	IleSerLysGluValGlyGluAlaSerIleLysValProGlnLeuLysMetGluIle	239
Db	1246	ATATCAAAAGAAAGTGGAGAAAGCTTCCATTAAGTACCAATTAATAAATGGAGATA	1302
RESULT 2			
US-10-067-443-23			
Sequence 23, Application US/10067443			
Patent No. 6642041			
GENERAL INFORMATION:			
APPLICANT: Bristol-Myers Squibb Company			
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE HIGHLY EXPRESSED IN			
FILE OF INVENTION: SPINAL CORD, MP-1			
FILE REFERENCE: D0073 NP			
CURRENT APPLICATION NUMBER: US/10/067,443			
CURRENT FILING DATE: 2002-02-05			
PRIOR APPLICATION NUMBER: US 60/266,518			
PRIOR FILING DATE: 2001-02-05			
PRIOR APPLICATION NUMBER: US 60/282,814			
PRIOR FILING DATE: 2001-04-10			
NUMBER OF SEQ ID NOS: 71			

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 1526
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-067-443-23

```

Alignment Scores:

1 150 157	1536
-----------	------

Score:	1240.00	Matches:	239
Percent Similarity:	100.00%	Conservative:	0
Best local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

```

Oy      1 LeuLeuAlaLeuValGInGlyValSerAspPheLeuLeuGlyLysSerLeuAsp1le 20
Db      85 CTGTTGGCATTAAGTTCAAGAGCTTCAGATTTTCCTCTTGGAAAGCTTTGGACATA 144
Oy      21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeu1lelyshisProGlu 40
Db      145 GCACCAAGGTGACATCTCTTGACAAGGTGGCAAGAAAGACTTTCTTAATAAATAACATCCAGAG 204
Oy      41 CysSerThrMetSerGlyGlyLysAla1leGluHisLeuAlaLysGInGlyAsnArgPhe 60
Db      205 TGCTTCACACATGAGGTGGTGGGAAAGCATTGAGAACATTGGCCAAACAAGAAATAAGATTT 264
Oy      61 HisPheAsp1leLysProProLeuHisshisAlaLysAsnGlyAspPheSerPheThrGly 80
Db      265 CATTGTGACATCAAACTCTTCCTTGCAATGCTAAATAATTTGATTTTTCTTTTACTGGA 324
Oy      81 LeuGlnHisValThrAspLys1le1leMetLysGlyLysGluGluGly1leGluLys 100
Db      325 CTTCACACGTTACTGATTAATAATAATGAAGAAAGAAAGAAAGCAAGGATTTGAGADAG 384
Oy      101 GlyGln1leLeuSerSerAlaAlaAsp1leAla1leAlaThrValGlnHisThrMetAlaCys 120
Db      385 GGGCAAAATCTCTCTTCACGACACAGACATTTGCTGCCACAGTACACACAAATGGCATGT 444
Oy      121 HisLeuValLysArgThrHisArgAla1leLeuPheCysLysGlnArgAspLeuLeuPro 140
Db      445 CATCTTGGAAGAAAGACACATCGGGCTATTTCTGTTTGTAGAGAGAGAGCTGTACT 504
Oy      141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPheTyr1leArgArg 160
Db      505 CAAATTAATGCGATCTGTTGCATCTGCGTGTGGCAAGTAACCTCTATATCCGACGA 564
Oy      161 AlaLeuGln1leLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLeu 180
Db      565 GCTCTGGAATTTTAAACAAACGCAACACAGTGCATTTGTTGTCTCTCCACAGACTA 624
Oy      181 CysThrAspAsnGly1leMet1leAla1leTrpAsnGly1leGluArgLeuArgAlaGlyLeu 200
Db      625 TGCACGTGATTAAGGAGATTAATGATTCATGATGAATGATTAAGAAAGACTACGCTGGCTTG 684
Oy      201 Gly1leLeuHisAsp1leGluGlyLysArgTyrGluProLysCysProLeuGlyValAsp 220
Db      685 GGCATTTTATCATGACATGAGAGGCAATCCGCTATGAACCAAAATGTCTCTTGGAGTGAC 744
Oy      221 HisSerLysGluValAlaGlyGluAlaSer1leLysValProGlnLeuLysMetGlu1le 239
Db      745 ATATCAAAAGAGGTGGAGAAAGCTTCATATAAAAGTCCACAATTAATAAATGAGATA 801

RESULT 3
US-10-067-443-1
; Sequence 1, Application US/10067443
; Patent No. 6642041
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNOCTLOBTIDE ENCODING A NOVEL METALLOPROTEASE HIGHLY EXPRESSED IN
; TITLE OF INVENTION: SPINAL CORD, MP-1

```

FILE REFERENCE: D0073 NP
CURRENT APPLICATION NUMBER: US/10/067,443
CURRENT FILING DATE: 2002-02-05
PRIOR APPLICATION NUMBER: US 60/266,518
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: US 60/282,814
PRIOR FILING DATE: 2001-04-10
NUMBER OF SEQ ID NOS: 71
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 2197
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (231)..(1472)
US-10-067-443-1

Alignment Scores:

Pred. No.:	2,07e-157	Length:	2197
Score:	1240.00	Matches:	239
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-649-273-2_COPY_176_414 (1-239) x US-10-067-443-1 (1-2197)

QY 1 LeuLeuAlaLeuValGInGlyValSerAspPheLeuLeuGlyLysSerLeuAsp11e 20
DB 756 CTGTGGCATTAGTTCAAGAGGTTTCAGATTTCGCTTTCGGAAGCTTTGGACATA 815
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 816 GCACCAAGGTACACGCTTGAACAGGTGCAAGAGACTTTTAAATAAACATCCAGAG 875
QY 41 CysSerThrMetSerGlyGlyValAlaIleGluHisLeuAlaLysGInGlyAsnArgPhe 60
DB 876 TGCTCCACCACTGAGTGGTGGAAAGCCATAGAACATTTGGCCAAACAGAAATAGATT 935
QY 61 HisPheAspIleLysProProLeuHisHisAlaLysAsnCyAspPheSerPheThrGly 80
DB 936 CATTTCACATCAAAACCTCCCTTCATCATCTAAAAATTGATTTTCTTTTACTGCA 995
QY 81 LeuGlnHisValThrAspLysIleIleMetLysLysGlyLysGluGluGly11eGluLys 100
DB 996 CTTCAACACCTTACTGATTAATATATATGAAAAAGAAAAAGAAAGTATTGACAG 1055
QY 101 GlyGlnIleLeuSerSerAlaAlaAspIleAlaIleThrValGlnHisThrMetAlaCys 120
DB 1056 GGGCAAAATCCGTCTTCAGCAGACATTCCTGCACAGTACAGACACAAATGCGATGT 1115
QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuPro 140
DB 1116 CATTTTGTGAAAAACACATCGGGCTATTCTGTTTGTAAAGCAGAGACTTGTAATCT 1175
QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaLysAsnPheThrIleArgArg 160
DB 1176 CAAATAATATGACATGACTGTTGATCTGGTGTGCGAAGTAATCTTCAATATCCGAGA 1235
QY 161 AlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
DB 1236 GCTCTGCAAAATTTTAAACAAAGCAACAGTGCATTTGTTGTGTCTCTCCACAGACTA 1295
QY 181 CysThrAspAsnGlyIleMetIleAlaThrAsnGlyIleGluArgLeuArgAlaGlyLeu 200
DB 1296 TGCACGTATATATGCAATTATATGCAATGATGATTAAGAAAGCTAGTGGCTTG 1355
QY 201 GlyIleLeuHisAspIleGluGly11eArgGlyArgProLysCysProLeuGlyValAsp 220
DB 1356 GGCATTTTACATGACATTAAGAGCATCCGCTATGAAACCAAAATGCTCTTGGAGTAGAC 1415
QY 221 IleSerLysGluValGlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 239

DB 1416 ATATCAAAAGAGTTGAGAGCTTCATTAAGTACCAATTAAATGAGATA 1472

RESULT 4

US-10-067-443-21

Sequence 21 Application US/10067443

Patent No. 6642041

GENERAL INFORMATION:

APPLICANT: Bristol-Myers Squibb Company

TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE HIGHLY EXPRESSED IN

TITLE OF INVENTION: SPINAL CORD, MP-1

FILE REFERENCE: D0073 NP

CURRENT APPLICATION NUMBER: US/10/067,443

CURRENT FILING DATE: 2002-02-05

PRIOR APPLICATION NUMBER: US 60/266,518

PRIOR FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: US 60/282,814

PRIOR FILING DATE: 2001-04-10

NUMBER OF SEQ ID NOS: 71

SOFTWARE: PatentIn version 3.0

SEQ ID NO 21

LENGTH: 1387

TYPE: DNA

ORGANISM: homo sapiens

US-10-067-443-21

Alignment Scores:

Pred. No.:	3.87e-154	Length:	1387
Score:	1213.50	Matches:	238
Percent Similarity:	90.15%	Conservative:	0
Best Local Similarity:	90.15%	Mismatches:	1
Query Match:	97.86%	Indels:	25
DB:	4	Gaps:	1

US-10-649-273-2_COPY_176_414 (1-239) x US-10-067-443-21 (1-1387)

QY 1 LeuLeuAlaLeuValGInGlyValSerAspPheLeuLeuGlyLysSerLeuAsp11e 20
DB 549 CTGTGGCATTAGTTCAAGAGGTTTCAGATTTCGCTTTCGGAAGCTTTGGACATA 608
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 609 GCACCAAGGTACATCTTGAACAGGTGCAAGAGACTTTTAAATAAACATCCAGAG 668
QY 41 CysSerThrMetSerGlyGlyValAlaIleGluHisLeuAlaLysGInGlyAsnArgPhe 60
DB 669 TGCTCCACCACTGAGTGGTGGAAAGCCATAGACATTTGGCCAAACAGAAATAGATT 728
QY 61 HisPheAspIleLysProProLeuHisHisAlaLysAsnCyAspPheSerPheThrGly 80
DB 729 CATTTCACATCAAAACCTCCCTTCATCATCTTAAAAATTGATTTTCTTTTACTGCA 788
QY 81 LeuGlnHisValThrAspLysIleIleMetLysLysGlyLysGluGluGly11e----- 98
DB 789 CTTCAACAGTACTGATTAATATATGAAAAAGAAAAAGAAAGTATTATTTCTTA 848
QY 98 ----- 98
DB 849 ATTACTAAAGTTGAACAGTAAATATCTCTGATGTGCTTAAATAATGCTGCTCATTTTC 908
QY 99 -----GluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaIleAlaThrValGln 115
DB 909 TGCAGGTATGAAAGGGGCAAAATCCGTCTTGAAGCAGACATTCCTCCACAGTACAG 968
QY 116 HisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGln 135
DB 969 CACACATGCGATGCTCATCTTCTGAAAGACACATCGGGCTATTCTGTTTGTAAAGCAG 1028
QY 136 ArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyValAsp11eSerAsn 155
DB 1029 AGAGACTTGTATCTCAAAATATATGACATGACTGTTGATCTGGTGTGCTCCCAAGTAGAC 1088
QY 156 PheThrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCys 175

Db 1089 TTTATATCCGAGAGCTCTGGAATTTTAAACCAACCAACAGCTGACCTTGTGTGT 1148
Qy 176 PROPRORPROARGHUCYSTRASPANGLYILEWETILEALATPASPANGLYILEGLUARG 195
Db 1149 CCTCTCCCACTAGCTAGCTGATTAATGCGATTATGATGATGATGATGATGATGATGAT 1208
Qy 196 LEUARGALAGLYLEUGLYILELEUHIASPILEGUIGLYILEARGTYGGLUPROLYCYE 215
Db 1209 CTACGATGATGCTTGGGCACTTTTACATGACATTAAGGCAATCCGCTATGAAACCAATGT 1268
Qy 216 PROLEUGLYVALSPILESERLYSGILUVALGILUVALASERILEYVALPROGLILEU 235
Db 1269 CCTCTTGAGATGACATATCAAAAAGAGTTGAGAAAGCTTCCATTAAGATCACAATTA 1328
Qy 236 LYMETGLULLE 239
Db 1329 AAAATGAGATA 1340
RESULT 5
US-10-067-443-20
; Sequence 20, Application US/10067443
; Patent No. 6642041
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE HIGHLY EXPRESSED I
; TITLE OF INVENTION: SPINAL CORD, MP-1
; FILE REFERENCE: D0073 NP
; CURRENT APPLICATION NUMBER: US/10/067,443
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 14364
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-067-443-20
Alignment Scores:
Pred. No.: 3 21e-86 Length: 14364
Score: 725.00 Matches: 186
Percent Similarity: 32.86% Conservative: 0
Best Local Similarity: 32.86% Mismatches: 0
Query Match: 58.47% Indels: 380
Gaps: 3
DB: 4
US-10-649-273-2_COPY_176_414 (1-239) x US-10-067-443-20 (1-14364)
Qy 29 VALAIAARGHUCYSTRASPANGLYILEWETILEALATPASPANGLYILEGLUARG 48
Db 11840 GTGGCAAGAGAGCTTTCTTTAATAAATCAACAGAGCTCCACATGAGTGGGAAA 11899
Qy 49 AIAIIEGLUHIASPILESERLYSGILUVALGILUVALASERILEYVALPROGLILEU 68
Db 11900 GGCATGAGACATTTGGCCAAACAGGAATGATTTGATTTGACATCAACCTCCCTTG 11959
Qy 69 HISHIALAVALASPILESERLYSGILUVALGILUVALASERILEYVALPROGLILEU 88
Db 11960 CATCATGCTAATAAATTTGATTTCTTTTACGAGCTCAACGTTACTGATTAATA 12019
Qy 89 ILEMETLYSGILUVALGILUVALASERILEYVALPROGLILEU 96
Db 12020 ATATATGAAAAAGAAAAAGAAAGATATTTCTATTAAGTGAACATTAAT 12079
Qy 97 -----GLYILEGLUVALGILUVALASERILEYVALPROGLILEU 103
Db 12080 ATTCCTGATTTGCTTAATAATAGTGTCTGATTTCTGACGATTTGAGAAAGGCAAT 12139
Qy 103 ELEUSERSERIALAASPILEALATMVALGIMHISTHRECALCYSHISLEUVA 123

Db 12140 CCTGTCTTCAGACACACACATCTGCGCACAGTACAGACCAATGCGATGCTTGTGT 12199
Qy 123 LIVESARGHUCYSTRASPANGLYILEWETILEALATPASPANGLYILEGLUARG 143
Db 12200 GAAAGAAACATCGGCTATTTGTTGTAAGACAGACATTTGTAACCAATAA 12259
Qy 143 PALAVALLEU----- 146
Db 12260 TGCAGTACTGTAAGTTTATCTCATTTATAGTAATAGTTACACTTGCATATGTAC 12319
Qy 146 ----- 146
Db 12320 TTTTTCACAGACCTTGACCTGTGTAGAGAACAGATCTTATGCTTATGCTAG 12379
Qy 146 ----- 146
Db 12380 CCTGACATGATGAATTTANGACGATAGAAAGCTAACGCCATTTCTGTACTAGTT 12439
Qy 146 ----- 146
Db 12440 TGATAGCTTATGAGACAGCTGTATAGCTTCTATGACATTAAGTCTAATTTGCACTT 12499
Qy 146 ----- 146
Db 12500 CTGTGTGATTTAAAGAGGCTTACATTAAGAAAGTAATGAGTAAGTCTATCACT 12559
Qy 146 ----- 146
Db 12560 ATTTTGAATAAATAGGTGATTTCTTCATCTTGTATGTAATCCCTTGTGTGTGT 12619
Qy 146 ----- 146
Db 12620 TTTTTAATAAGCAATCAATTTAGCAGTGGAGGTGTATTCACACTTGTGACACT 12679
Qy 146 ----- 146
Db 12680 AATGTGATAAAGTCTGATATCACTATATTTGACAGCCAAATCCCTTAATGTG 12739
Qy 146 ----- 146
Db 12740 CTTAAAGCTTGACAAACATCTGTATTAAGCTTAACTTATCAATTTAAAT 12799
Qy 146 ----- 146
Db 12800 TATTAACATAAGTGGAAATGTTAATGCTAGTAATCATAGATGAATTTACATGG 12859
Qy 146 ----- 146
Db 12860 ATATCAAGAAATATTTTTCAGAGTTATGTAGTAATAAGCAAAATATTAATAATTC 12919
Qy 146 ----- 146
Db 12920 AGGCTTAATAATAGTACTATGATTAATTAATTAATTAATTAATTAATTAATTAAT 12979
Qy 146 ----- 146
Db 12980 TGAAGAAATATTAATAAATGCTAGTAATGTTGTATGCTAATTAATTAATTAAT 13039
Qy 146 ----- 146
Db 13040 TTTTCTTCCAAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 13099
Qy 146 ----- 146
Db 13100 GGATAGTTATTAATTTGTTAATGCTAATTTTTCAGAGTTAATTAAGAGCTTGT 13159
Qy 147 -----VALAASERILEYVALPROGLILEU 151
Db 13160 TCATATCATTAATAGTATTTGTTTCTCAATCTTCAAGTTGATCTGAGTCTGAGTGG 13219
Qy 151 YVALAASERINPHELYILEARGALALEUGLYILEUHIASPILEALATMVALGIMHISTHRECALCYSHISLEUVA 171

DB 13220 TGTGCAAGTAAGTCTTATATCCGACAGCTCTGAAATTTTAAACCAACGACAGCTG 13279
QY 171 sThrLeuLeuCyseProProProProLeuCyseThraAspAsnGlyIleMetIleAlaTrp-- 190
DB 13280 CACTTTGTTGTCTCTCTCCAGACATATGACATGATGATGATGATGATGATGATGATGAT 13339
QY 190 ----- 190
DB 13340 AGCCACAGATATATGCTTCTCACTCATTAATATATTAATGATGATGATGATGATGAT 13399
QY 190 ----- 190
DB 13400 ACTAAGCTTCTCTCTCTCACTCATTAATATATTAATGATGATGATGATGATGATGAT 13459
QY 191 --AsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleA 210
DB 13460 GGAATGATATGAAAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13519
QY 210 rGTYrGluProlys 214
DB 13520 GCTATGACCAAG 13533
RESULT 6
US-09-596-002-38
Sequence 38, Application US/09596002
Patent No. 6632636
GENERAL INFORMATION:
APPLICANT: Lagace, Robert, E.
APPLICANT: Patterson, Chandra
APPLICANT: Berg, Klm, E.
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
FILE REFERENCE: PM-0008-4 US
CURRENT APPLICATION NUMBER: US/09/596,002
PRIOR FILING DATE: 2000-06-16
PRIOR FILING DATE: 60/140,121
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PERL Program
SEQ ID NO 38
LENGTH: 94750
TYPE: DNA
ORGANISM: M. catarrhalis
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: incyte template ID No. 6632636 38
US-09-596-002-38
Alignment Scores:
Pred. No.: 3,02e-20 Length: 94750
Score: 248.50 Matches: 73
Percent Similarity: 48.67% Conservative: 37
Best Local Similarity: 32.30% Mismatches: 90
Query Match: 20.04% Indels: 27
Gaps: 7
DB: 4
US-10-649-273-2_COPY_176_414 (1-239) x US-09-596-002-38 (1-94750)
QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLeuSerLeuAspIle 20
DB 24305 ATGCTGTGCGTGGCGGATGCTGTGGCGGTGATCAGATATTTGGGAGATCTATCATGAT 24364
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 24365 GCGGTGGGGAATGCTTGAATAAACGCAAAATGCTCAATG---CCCTATCTCT--- 24418
QY 41 CysSerThrMetSerGlyGlyLysAlaIleGluHisLeuAlaLysGlnGlyAsnArgPhe 60
DB 24419 -----GGTGGCGCCAAATATGCAAAATTTAGCCAAATTAACGCAACGCAAC 24463
QY 61 HisPheAspIleLysProProLeuHisAlaLysAsnCyseAspPheSerPheThrGly 80
DB 24464 GCGTATAGCTGCCAAGACCATGACGAT---AAAGGCTGAGATTTTTCCTTCAGTGC 24520

QY 81 LeuGlnHisValThrAspLysIleMetLysLysGlyLysGluGluGlyIleGluLys 100
DB 24521 ATGAAACCGCATTCATTAATCTCAAGACACACAAACGCAACGCAACGCAACGCAAC 24577
QY 101 GlyGlnIleLeuSerSerAlaLysAspIleAlaIleAlaThrValGlnHisThrMetAlaCys 120
DB 24578 -----GCCACAGACAGACATCGCGCAAGCTTGAATATGCGGTGGAT 24625
QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
DB 24626 ACTTGGCTCAAAAATGACCAACACTACAGATGACAGGATTCGCCAG----- 24676
QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyValAlaSerAspPheThrIleArgArg 160
DB 24677 -----CTGGTGTCCAGAGGGGCGTCTGCCAATCAGATGCTACGCCCC 24721
QY 161 AlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCyseProProArgLeu 180
DB 24722 ACCCTGACCGAGCCCTCCGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 24781
QY 181 CysThrAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeu 200
DB 24782 TGCACGATATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 24841
QY 201 GlyIleLeuHisAspIleGluGlyIleArgGlyGluProLysCyseProLeuGlyValAsp 220
DB 24842 TCG-----GATGATTTGGC-GGTTGCGGTATTC-----CCGATGGATATATGC 24885
QY 221 IleSerLysGluValGly 226
DB 24886 GACGCTTGCGGATCGGC 24903
RESULT 7
US-09-540-236-806
Sequence 806, Application US/09540236
Patent No. 6673910
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
FILE REFERENCE: 2709,2005-001
CURRENT APPLICATION NUMBER: US/09/540,236
PRIOR FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO 806
LENGTH: 1053
TYPE: DNA
ORGANISM: M. catarrhalis
US-09-540-236-806
Alignment Scores:
Pred. No.: 3,22e-23 Length: 1053
Score: 247.00 Matches: 70
Percent Similarity: 46.85% Conservative: 34
Best Local Similarity: 31.53% Mismatches: 80
Query Match: 19.92% Indels: 38
Gaps: 6
DB: 4
US-10-649-273-2_COPY_176_414 (1-239) x US-09-540-236-806 (1-1053)
QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
DB 445 ATGCTGTGCGTGGCGGATGCTGTGGCGGTGATCAGATATTTGGGAGATCTATCATGAT 504
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 505 GCGGTGGGGAATGCTTGAATAAACGCAAAATGCTCAATG---CCCTATCTCT--- 558
QY 41 CysSerThrMetSerGlyGlyLysAlaIleGluHisLeuAlaLysGlnGlyAsnArgPhe 60
DB 559 -----GCTGTCCAAATATGCAAAATTTAGCCAAATTAACGCAACGCAACGCAAC 603

QY		61	HISPLEADPIELYSProProblemishialalyasyencyasppheSerPheThcIly	80
Dd		604	GCTATGAGCTGCCAAGCCCATGTACAT---AAAGGGCTGAGATTTCCTGTACGTGC	660
QY		81	LeuGlInhiSValThrAspLysilelleMetLysylsGiLysglLuglYilleLlys	100
Dd		661	ATGAANAACCGCATTCAATATCTCATCAAAGAACAAGCCCAAGGCCCAAGGAGACCT---	717
QY		101	GlycInulleuSerSerAlalaspIlealalathrValGIlnhiStHmetAlaCys	120
Dd		718	-----GCCACACAGCAGACATCCGCCCAACTTTTGAGTATGACGGTGTGAT	765
QY		121	HisleuVallysArgThrhisArgAlalleuPheCysLysGIlnArgAspleuLeuPro	140
Dd		766	ACTTTGGTGCAAAAATGCAACAGCGCATCATGATGACAGGCATTTCGGCAG-----	816
QY		141	GIlnASnaSnalAvalleuValAlaserGIgLyValAlaserAsnPhetyrIleArgArg	160
Dd		817	-----CTGTGTGTGCGAAGGGGGCGTCTTGCCAATCAGACGCTACGGCGC	861
QY		161	AlaleuGlInulleuThrAsnalArthrGIncysThrleuLeuCysProProArgIreu	180
Dd		862	AACCTGACCGAGAGCGCTCGGCCAATGCATGCGTGCGTACTACGCCCCACAGCACTA	921
QY		181	CysThrAspaanglylleMetIleAlaTrpasnglylleGIunArgleuArgAlagly---	199
Dd		922	TGCACGGATATATGGTCGATGATCGCTTAACCTGGCTTTGTCCGCTTAAGCCGTGGCAG	981
QY		200	-----LeuGIYilleu	203
Dd		982	TGGATGACTTGGCGGTTCGCTGATATCCCGATGGATATGACGATGACGATCGTTGTATCGAA	1041
QY		204	HISASP	205
Dd		1042	TATGAT	1047

US-09-252-991A-884/c
Sequence 884, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 884
LENGTH: 1059
TYPE: DNA
ORGANISM: *Pseudomonas aeruginosa*
US-09-252-991A-884

Alignment Scores:		
Pred. No.:	4,81e-21	1059
Score:	231.00	Matches:
Percent Similarity:	49.54%	Conservative:
Best Local Similarity:	33.80%	Mismatches:
Query Match:	18.63%	Indels:
DB:	4	Gaps:

Qy 2 LeuAlaLeuValGlnGlyValSerPheLeuLeuGlyLysSerLeuAspIleAa 21
 ||| ||| :: ||||| :||| |||
Db 621 TTGGTGCAGGTGAACGGATTCGACCCTCAACTCTTGAGCAATCGGTGACAGATGCC 562
 ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 22 ProGlyAspMetLeuAspLysValAlaAlaGluSerLeuLelyShiSPProGluCys 41

Db 561 GCGCGGAGGCGCTTGACAGAGCGCGGCGAGCGATGCGGCTG---GGCTATGCC-----5111

Gy 42 SerThmetSerGlyGlyLysalaaleguintsleuallayeglnllyAsnargphenis.61

Db 510 -----GGTGTGTCGGAAATCGCGCGCTGCGGAGGCGGCACTCTGCGCGC4633

Gy 62 PheAspIleLysProProLeuHisHisalaLysAspCysAspPheSerPheThrGlyLeu81

Db 462 TTTCGGTTCGCGCGCGCGCGATCGCGCGCGCTGAGCTTCACGTTTCACCGGAGCTC4033

Gy 82 GlnHisValThrAspLysIleLeuMetLysLysGlyLysGlnGlnGlyIleGlnLysGly101

Db 402 AAGACCTTTACCCTGACAC--ACCTGGCGAGCGCTGCGAGAGCGCGGCGACGCGAG346

Gy 102 GlnIleLeuSerSerSerLalaAspIleAlaIleValGlnHisIleThrMetAlaCysHis121

Db 345 CAG-----ACCGCGCTGCAGACATCGCGCTTCGAGCGCGCGGCGTGTGAGACC2955

Gy 122 LeuValLysArgThrHisIserGlnAlaIleLeuPheCysLysGlnArgAspLeuProGln141

Db 294 CTGCTGATCAAGTGGCGGCTGCGGCTTG-----AACCAAGCGCGCTG-----AAG250

Gy 142 AsnAspAlaValLeuValAlaSerGlyLysValAlaSerAspPheThrIleArgArgAla161

Db 249 AAC-----CTGTGATGCGCGCGGCTGTGAGGCGACACGAGCGCTGCGCAGCGCG199

Gy 162 LeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProAlaGlyLeuCys181

Db 198 CTGGAAAGATGCTCGCGCGAAATGGAAGGGGAGAGTGTCTACAGCGCGCGCGCTTCTGC139

Gy 182 ThrAspAsnGlyIleMetIleAlaThrAsnGlyIleGlnArgLeuArgAlaGlyLeuGly201

Db 138 ACCGACATGGGCGCATGATCGCTCAAGCGCGGCTGCGACGCGCTGCTGCGCGGCG-----85

Gy 202 IleuHisAspIleGlnGlyIleArgThrGlnProLysCysProLeu217

Db 84 ---CAGCATGTACGCGCGCGCGATGACAGCTTCAGCGCGCTGCGCGATG40

RESULT 9
 US-09-252-991A-801
 : Sequence 801, Application US/09252991A
 : Patent No. 6551795
 : GENERAL INFORMATION:
 : APPLICANT: Marc J. Rubenfield et al.
 : TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 : TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 : FILE REFERENCE: 107196.136
 : CURRENT APPLICATION NUMBER: US/09/252,991A
 : CURRENT FILING DATE: 1999-02-18
 : PRIOR APPLICATION NUMBER: US 60/074,788
 : PRIOR FILING DATE: 1998-02-18
 : PRIOR APPLICATION NUMBER: US 60/094,190
 : PRIOR FILING DATE: 1998-07-27
 : NUMBER OF SEQ ID NOS: 33142
 : SEQ ID NO 801
 : LENGTH: 1206
 : TYPE: DNA
 : ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-801

Alignment Scores:	
Pred. No.:	5, 95e-21
Score:	211.00
Percent Similarity:	49.54%
Best Local Similarity:	33.80%
Query Match:	18.63%
DB:	4
Gaps:	8
Length:	1206
Matches:	73
Conservative:	34
Mismatches:	87
Indels:	22

```

US-10-649-273-2_COPY_176_414 (1-229) x US-09-252-991A-801 (1-1206)
Qy      2  LeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlnGlyysSerLeuAspIleVal 21
      |||  |||  |||:::  ::|  |||||:::  |||:::  |||  |||

```



```

Db 604 TTGTCGGGTCGACGGTATCGCGCTACGACCTTGGCGAATCGTCGACGATGCC 663
Qy 22 ProGlyAspMetLeuAspLysValAlaArgLysSerLeuIleuHisProGlyCys 41
Db 664 GCGGGGAAGCCTTCCAGACAGACCGCAGACCTATCGGCTG---GGCTATCCC----- 714
Qy 42 SerThrMetSerGlyGlyLysAlaIleGluHisLeuAlaLysGlnGlyAsnArgPheHis 61
Db 715 -----GGTGGTCGGGAAATCGCGGCTGGCGGAGCGCGGCACTCCGCGCGC 762
Qy 62 PheAspIleLysProLeuHisHisAlaLysAsnCysAspPheSerPheThrGlyLeu 81
Db 763 TTGCGTGTTCGGCGGCGATCGACCGATCGCGCGCTCCGACCTTCAGCGGCTC 822
Qy 82 GlnHisValThrAspLysIleIleMetLysGlnLysGlnGlyIleGluLysGly 101
Db 823 AAGACCTTACCTTGAC---ACCTGGCAGCGTTGCTTCAGAGCGGCGAGACAGCGAG 879
Qy 102 GlnIleLysSerSerAlaIleAspIleAlaIleThrValGlnHisThrMetAlaCysHis 121
Db 880 CAG-----ACCGCTCGGACATCGCCCTGCGCTCCAGACCGCGGCTGTGACAGC 930
Qy 122 LeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuProGln 141
Db 931 CTGCTGATCAGTGCCTGCGGCTTG-----AAGCAGACCGGCTG-----AAG 975
Qy 142 AsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerPhePheThrIleArgAla 161
Db 976 AAC-----CTGCTGATCGCGCGGCGGTGTCAGCCCAACAGCGGCTGCGAGCGC 1026
Qy 162 LeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLysCys 181
Db 1027 CTGGAAGAAGATGTCGCGCAATGAAGGCGCAGGTGTTTACGCCCGCGCTTCTGC 1086
Qy 182 ThrAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeuGly 201
Db 1087 ACCGACAAATGGCGGATGATCGCTTACGCCGCGTCCAGCGCGCTGCTGCGCGC----- 1140
Qy 202 IleLeuHisAspIleGluGlyIleArgGlyLeuGluProLysCysProLeu 217
Db 1141 ---CAGCATGACGCGCGCGCATCGAGTCCAGCGCGCTGCGCGCATG 1185

```

RESULT 10
 US-09-557-884-1/c
 ; Sequence 1, Application US/09557884
 ; Patent No. 6506581
 GENERAL INFORMATION:
 APPLICANT: Fieischmann et al.
 TITLE OF INVENTION: The Nucleotide sequence of
 the Haemophilus influenzae Rd Genome, Fragments
 thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: MD
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3 1/2 inch diskette
 COMPUTER: Dell Pentium
 OPERATING SYSTEM: MS DOS V6.22
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/557,884
 FILING DATE: 25-Apr-2000
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/476,102
 FILING DATE: JUN-5-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971
 REFERENCE/DOCKET NUMBER: P8186F3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 301-309-8504
 TELEFAX: 301-309-8439
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1830121 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-557-884-1

Alignment Scores:
 Pred. No.: 7,86-15 Length: 1830121
 Score: 224.00 Matches: 67
 Percent Similarity: 47.50% Conservative: 28
 Best Local Similarity: 33.50% Mismatches: 83
 Query Match: 18.06% Indels: 22
 DB: 4 Gaps: 6

US-10-649-273-2_COPY_176_414 (1-239) x US-09-557-884-1 (1-1830121)

```

Qy 2 LeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIleAla 21
Db 552791 TTACTGGGCTGTCGATGGGTGATGAAATATGAAATGATAGAGATGATATGATGCT 552732
Qy 22 ProGlyAspMetLeuAspLysValAlaArgLysSerLeuIleuHisProGlyCys 41
Db 552731 GCTGGCAGAGCCTTGTGATTAACAGCAAAATTACTTGACCTA---GATTATCCA----- 552681
Qy 42 SerThrMetSerGlyGlyLysAlaIleGluHisLeuAlaLysGlnGlyAsnArgPheHis 61
Db 552680 -----GGTGGCGCGGCACTTCTGCTTAGCGGAAAGGTAACCCCAATGCT 552633
Qy 62 PheAspIleLysProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGlyLeu 81
Db 552632 TTTCATTTCCACGTCATATACAGATCGTCAGCGCTTGAATTTTTCGGTTA 552573
Qy 82 GlnHisValThrAspLysIleIleMetLysGlnLysGlnGlyIleGlu 99
Db 552572 AAAACATTGGCGGCAATATACGTTAATCAACAAATTAAGGAGGCGCAATGATAG 552513
Qy 100 LysGlyGlnIleLysSerSerAlaIleAspIleAlaIleThrValGlnHisThrMetAla 119
Db 552512 CAA-----ACTAAAGCAGATATGCTTATGCTTTCAGAGATGCGGTGCTG 552468
Qy 120 CysHisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeu 139
Db 552467 GATACTCTTGCC-----ATTAAATGTAG--CGTGCATTGAAA 552432
Qy 140 ProGlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPheThrIleArg 159
Db 552431 GAAACAGCTATAACGTTTATGATTCGCGGAGGCGTGAAGCGCAATTAAGAACTCCGA 552372
Qy 160 ArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArg 179
Db 552371 GAAACGCTTGGCAGCTATATGCAAAATTTAGGTGGAGGATTTTATCTCAACTCAA 552312
Qy 180 LeuCysThrAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuArgAlaGly 199
Db 552311 TTTTGTACAGATTAATGTGCGCATGATGCTTACACAGGTTTTTTAAACAGAGGT 552252

```

RESULT 11
 US-09-643-990A-1/c
 ; Sequence 1, Application US/09643990A
 ; Patent No. 6528289
 GENERAL INFORMATION:
 APPLICANT: Robert D. Fieischmann
 Mark D. Adams
 Owen White
 Hamilton O. Smith

Db 685 GGGCAGCGCGCTG-----GGGATTTGTGCGCTCTTCCAGAGCGCCGTGCGGAC 735
Qy 121 HisLeuValIysArgThrHisArgAlaIleLeuPheCysIysGlnArgAspLeuPro 140
Db 736 GTGCTGTCAAGAG-----CTGTGCGCCCGCGCGCGCGGTG----- 774
Qy 141 GlnAsnAlaValIleuValAlaSerGlyGlyValAlaSerAspPheTyrIleArgArg 160
Db 775 ---GGCCACAGAGATTGTGTGTGTGCGCGCGCGCGCGGACCTCGCGGCTCGGCA 831
Qy 161 AlaLeuGluIleLeuThrAsnAlaThrGlnCysThrIleLeuCysProProProArgLeu 180
Db 832 CTGTGTCAAGCGCGAGCGCGAGCGCGGCGTTGAACATGTTCTGCGCCCGGTGCGCTG 891
Qy 181 CysThrAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeu 200
Db 892 TGCACGACATATGCGCCCATATGATGCGGTGCGCGGCGGTATAGCGCGTACCGCGCGGCTG 951

RESULT 13
US-09-902-540-503/C
Sequence 503, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 503
LENGTH: 2582
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-503

Alignment Scores:
Pred. No.: 3,98e-19 Length: 2582
Score: 221.50 Matches: 70
Percent Similarity: 47.50% Conservative: 25
Best Local Similarity: 35.00% Mismatches: 74
Query Match: 17.86% Indels: 31
Gaps: 8

US-10-649-273-2_COPY_176_414 (1-239) x US-09-902-540-503 (1-2582)

Qy 5 ValGlnGlyValSerAspPheLeuLeuGlyIysSerIleAlaIleProGlyAsp 24
Db 1096 GTCAAGCCCTACAGCGGAGTACCGGCTGTGTGGAGACAGCGGACAGACGCGCGCGGAG 1037
Qy 25 MetLeuAspIysValAlaArgArgLeuSerIleuIleuIleuIleuIleuIleuIleuIleu 44
Db 1036 GCATATACAAAGCCGCTCGCATCTCGGCTG---CCGATACCG----- 995
Qy 45 SerGlyGlyValAlaIleGlnIleuAlaIysGlnIysAsn-----ArgPhe 60
Db 994 ---GGTGGAGCGCCATGACAGTGGCGAGAGGGGAAACCGGAGCGCATCCGCTTC 938
Qy 61 HisPheAspIleuIleuProIleuHisIleuAlaIysAsnCysAspPheSerPheThrGly 80
Db 937 -----CCGCGCGCGCTCGCGCGCGAGCAACTTCGATCTCTTCCTCGCGG 893
Qy 81 LeuGlnHisValThrAspIleIleMetIysIysGlyIysGlyGlyIleGlyIys 100
Db 892 TTGAAG-----ACGCGGTGTGTGACCACTGACAGAGCGGCGTGCAG 845
Qy 101 GlyIleLeuSerSerAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 120
Db 844 GGGCAGCGCGCTG-----GCGGATTTGTGCGCGCTCTTCCAGAGCGCGGTGCGGAC 794

Qy 121 HisLeuValIysArgThrHisArgAlaIleLeuPheCysIysGlnArgAspLeuPro 140
Db 793 GTGCTGTCAAGAG-----CTGTGCGCCCGCGCGCGCGGTG----- 755
Qy 141 GlnAsnAlaValIleuValAlaSerGlyGlyValAlaSerAspPheTyrIleArgArg 160
Db 754 ---GGCCACAGAGATTGTGTGTGTGCGCGCGCGCGCGGACCTCGCGGCTCGGCA 698
Qy 161 AlaLeuGluIleLeuThrAsnAlaThrGlnCysThrIleLeuCysProProProArgLeu 180
Db 697 CTGTGTCAAGCGCGAGCGCGAGCGCGGCGTTGAACATGTTCTGCGCCCGGTGCGCTG 638
Qy 181 CysThrAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeu 200
Db 637 TGCACGACATATGCGCCCATATGATGCGGTGCGCGGCGGTATAGCGCGTACCGCGCGGCTG 578

RESULT 14
US-09-543-681A-2341
Sequence 2341, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 2341
LENGTH: 1074
TYPE: DNA
ORGANISM: Proteus mirabilis
US-09-543-681A-2341

Alignment Scores:
Pred. No.: 1.86e-18 Length: 1074
Score: 212.00 Matches: 66
Percent Similarity: 44.95% Conservative: 32
Best Local Similarity: 30.28% Mismatches: 90
Query Match: 17.10% Indels: 30
Gaps: 7

US-10-649-273-2_COPY_176_414 (1-239) x US-09-543-681A-2341 (1-1074)

Qy 2 LeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyIysSerIleAlaIleProGlyAsp 21
Db 475 TTAAATTAGTGAACAGGATTCGGGAATATACCTCTTAGTGAAGTATGATGATGCT 534
Qy 22 ProGlyAspMetLeuAspIysValAlaArgArgLeuSerIleuIleuIleuIleuIleuIleuIleu 41
Db 535 GGTGTGAAGCATTTGATTAACAGCAAGCATATGCGGCTT---GATTAATCC----- 585
Qy 42 SerThrMetSerGlyIysValAlaIleGlnIleuAlaIysGlnIysAsnArgPheHis 61
Db 586 -----GGGCGCGCTGTATTATCAAAATGCGACAAAGAGTGTAGAGAGACT 633
Qy 62 PheAspIleuIleuProIleuHisIleuAlaIysAsnCysAspPheSerPheThrGlyLeu 81
Db 634 TTGTGTTTCTCGTCCATGACAGACAGACCGGAGCTGACCTTTAGTTGTTCTGTGTTA 693
Qy 82 GlnHisValThrAspIleIleMetIysIysGlyIysGlyGlyIleGlyIysGly 101
Db 694 AAACCTTGGCGCGCATATGATTCGTAACAAAGATGATTCAGAG----- 738
Qy 102 GlnIleLeuSerSerAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 121
Db 739 -----CAAACTGACAGATATGCGCGCTGTTGAAGAGCCGCTAGTACT 789
Qy 122 LeuValIysArgThrHisArgAlaIleLeuPheCysIysGlnArgAspLeuProGln 141

Db 790 TTGGCAATAAATGCTGACGA-----TTAGACAA 822
Qy 142 AsnAsn---AlaValLeuValAlaSerGlyGlyValAlaSerAsnPhetYrIleArg 160
Db 823 ACAGCGCTTAAACGCTTAGATGATGCTGGGGCGGTAAAGCTTAACGCGCC 882
Qy 161 AlaleuGluIleleuThrAsnAlaThrGlnCysThrIleuLeuCysProProProArgLeu 180
Db 883 AAAATGCGCATGTATGAAACAACCTCGAGGGGAAGTGTATGCTCGCCCTGAGCTTA 942
Qy 181 CysThrAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeu 200
Db 943 TGTACCGATAATGATGCTCATGATGCTTGGCGGGAGATGCCGTTTAAAGGTGTAAC 1002
Qy 201 ---GlyIleleuHisAspIleGluGlyIleArgTyrgIuProlysCysProleu 217
Db 1003 GAGGGGCGCTTTA-----GGGGTGACAGTGAAGACCACTGGCCTTTA 1044

RESULT 15

US-09-489-039A-2050
; Sequence 2050, Application US/09489039A
; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA; FILE REFERENCE: 2709.2004001
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/489,039A

; PRIOR FILING DATE: 2000-01-27

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 2050

; LENGTH: 1032

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-2050

Alignment Scores:
Pred. No.: 1.55e-17 Length: 1032
Score: 205.00 Matches: 66
Percent Similarity: 43.58% Conservative: 29
Best Local Similarity: 30.28% Mismatches: 93
Query Match: 16.53% Indels: 30
DB: 4 Gaps: 7

US-10-649-273-2_COPY_176_414 (1-239) x US-09-489-039A-2050 (1-1032)

Qy 2 LeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIleAla 21
Db 442 TTGATTAGCGGTACCGGATATGCTAGTAACTGCTGGCGAGTGAAGATGCG 501
Qy 22 ProGlyAspMetLeuAspIleValAlaArgArgLeuSerLeuIleLysProGlnCys 41
Db 502 GCGGGGAGCGCTTATATAGACGCGAAGCTCTGGACTG--GATTATCCC----- 552
Qy 42 SerThrMetSerGlyLysAlaIleGlnHisLeuAlaLysGlnGlyAsnArgPheHis 61
Db 553 -----GCGCGGCGGATGCTGCGAAAATGCGTCCGAGGCGACCGAAGCGCGC 600
Qy 62 PheAspIleLysProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGlyLeu 81
Db 601 TTGTGTTTCCCGGCGGATGACCGACCGGCGGCTGACATTCACTTCCGCGCTG 660
Qy 82 GlnHisValThrAspLysIleIleMetLysLysGlnLysGlnGlyIleGluLysGly 101
Db 661 AAGACCTTCGCGCAACACATTCCAGAACGCGACGACGATGAG-----705
Qy 102 GlnIleLeuSerSerAlaAspIleAlaIleAlaThrValGlnHisThrMetAlaCysHis 121
Db 706 -----CAAAACCGCGCGACATCCCGCGCGCTTTAGAGATGCGGTGCGATACG 756
Qy 122 LeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuProGln 141

Db 757 CTGATGATTAAATGCTCGCGGCG-----CTGAGACAA 789
Qy 142 AsnAsn---AlaValLeuValAlaSerGlyGlyValAlaSerAsnPhetYrIleArg 160
Db 790 ACCGCGCTTAAACGCTTAGATGATGCTGGGGCGGTAAAGCTTAACGCGCC 849
Qy 161 AlaleuGluIleleuThrAsnAlaThrGlnCysThrIleuLeuCysProProProArgLeu 180
Db 850 AAGCTGCGGAGATGATGACAAAACGCGCGCGAGGTGTCTAGCCCGCCTGAGTTC 909
Qy 181 CysThrAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeu 200
Db 910 TGTACTGACAAACGCGGATGATGCTCATGCGCGGATGCTGCTGCTGCAAAACGCGCGCC 969
Qy 201 GlyIleleuHisAspIleGlu---GlyIleArgTyrgIuProlysCysProleu 217
Db 970 -----AAAGCGAGCTGCGGCTGACGCTTGGCCCGCGCTGCGCGCTG 1011

Search completed: June 17, 2005, 02:30:33
Job time : 984.721 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 16, 2005, 23:18:53 ; Search time 1623.64 Seconds

(without alignments)
913.751 Million cell updates/sec

Title: US-10-649-273-2_COPY_176_414

Perfect score: 1240
Sequence: 1 L1ATQGVSDFLGLKSLDI.....DISKEVGASIKVPQKMEI 239

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6054689 segs, 3103772919 residues

Total number of hits satisfying chosen parameters: 12109378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_plus.p2n.model -DEV=slp
-O=/cgn2_1/USPTO.spool.p/US10649273/runat_15062005_111418_6138/app.query.fasta_1.1429
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=p2n.rmp -MINMATCH=0.1
-LOOPT=0 -LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USR=US10649273_@CGN_1_1_1034_@runat_15062005_111418_6138
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WMIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US10J_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1240	100.0	1416	17	US-10-120-988-177
2	1240	100.0	1526	14	US-10-067-443-23
3	1240	100.0	1526	18	US-10-649-273-23
4	1240	100.0	1526	18	US-10-651-722-23
5	1240	100.0	2197	14	US-10-067-443-1
6	1240	100.0	2197	18	US-10-649-273-1
7	1240	100.0	2197	18	US-10-651-722-1
8	1240	100.0	2572	21	US-10-480-988-36
9	1233.5	97.9	1387	18	US-10-649-273-21
10	1233.5	97.9	1387	18	US-10-651-722-21
11	1233.5	97.9	1387	21	US-10-887-553A-1047
12	1233.5	97.9	1387	21	US-10-887-553A-1047
13	1203	97.0	1245	14	US-10-012-140-6
14	1203	97.0	1820	14	US-10-012-140-4
15	1059	85.4	2208	17	US-10-094-749-400
16	1059	85.4	2890	20	US-10-723-860-7447
17	725	58.5	14364	14	US-10-067-443-20
18	725	58.5	14364	18	US-10-649-273-20
19	725	58.5	14364	18	US-10-651-722-20
20	468	37.7	371	19	US-10-430-201-3118
21	468	37.7	371	19	US-10-430-201-3118
22	337	27.2	1917	18	US-10-424-599-66417
23	270	21.8	1628	19	US-10-437-963-11249
24	257	20.7	1146	17	US-10-282-122A-14674
25	248.5	20.0	1000	19	US-10-343-561-50
26	248.5	20.0	1044	17	US-10-282-122A-26972
27	248.5	20.0	94750	18	US-10-672-787-38
28	247	19.9	936	17	US-10-282-122A-8315
29	241	19.4	1032	17	US-10-282-122A-31043
30	239	19.3	756	14	US-10-081-051-8
31	239	19.3	4360	14	US-10-081-051-2
32	231	18.6	1026	9	US-09-815-242-7701
33	231	18.6	1026	17	US-10-282-122A-30016
34	224	18.1	1029	9	US-09-815-242-6946
35	224	18.1	1029	17	US-10-282-122A-22020
36	224	18.1	1830121	17	US-10-329-670-1
37	224	18.1	1830121	20	US-10-158-865-1
38	213	17.2	1014	9	US-09-815-242-9682
39	213	17.2	1014	17	US-10-282-122A-39301
40	212	17.1	1020	17	US-10-282-122A-32254
41	210	16.9	1014	17	US-10-282-122A-41977
42	208	16.8	1007	17	US-10-282-122A-19320
43	208	16.8	1728	19	US-10-437-963-9137
44	207.5	16.7	1023	17	US-10-282-122A-31809
45	206	16.6	1014	9	US-09-815-242-6207

ALIGNMENTS

RESULT 1
US-10-120-988-177
Sequence 177, Application US/10120988
Publication No. US20030219745A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Goodrich, Ryle
APPLICANT: Liu, Chenghua
APPLICANT: Ren, Feiyun
APPLICANT: Wang, Dunrui
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. US20030219745A1 Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 802CON
CURRENT APPLICATION NUMBER: US/10/120,988
CURRENT FILING DATE: 2002-04-11

PRIOR APPLICATION NUMBER: 09/774,528
PRIOR FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 441
SOFTWARE: pc_Fl_genes Version 2.0
SEQ ID NO 177
LENGTH: 1416
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (205)..(1305)
US-10-120-988-177

Alignment Scores:

Pred. No.:	1,476-151	Length:	1416
Score:	1240.00	Matches:	239
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	17	Gaps:	0

US-10-649-273-2_COPY_176_414 (1-239) x US-10-120-988-177 (1-1416)

QY 1 LeuLeuAlaLeuValGInGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
DB CTGTTGGCATTAGTTCAGAGAGATTTCAGATTTCCTGCTTGGAAAGTCTTGGACATA 645
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuLeuIleLysIleProGlu 40
DB GACCCAGGTGACATGCTTGAAGAAGTGGCAAGAGACCTTCTTAAATAAACATCCAGAG 705
QY 41 CysSerThrMetSerGlyLysAlaIleGluHisLeuAlaLysGlnGlyAsnArgPhe 60
DB TGTCTCCACCATGAGTGGGAAAAGCCATGAAACATTTGGCCAAACAGAAATAGATT 765
QY 61 HisPheAspIleLysProProLeuHisIleAlaLysAsnCysAspPheSerPheThrGly 80
DB CATTTGACATCAAACTCCCTTCATCATGCTAAATAATGCTATTTCTTTTACTGGA 825
QY 81 LeuGlnHisValThrAspLysIleIleMetLysLysGluLysGluGluGlyIleGluLys 100
DB CTTCAACACCTTCTGATTAATAATATATGAAAAAGAAAAAGAGATTGACAAAG 885
QY 101 GlnGlnIleLeuSerSerAlaIleAspIleAlaIleArgValGlnHisIleMetAlaCys 120
DB GGGCAAAATCTGCTTCAGCAGCAGACATGCTGCCACAGTACGACACACATGGCATGT 945
QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
DB CATCTTGTGAAAAGAACACATCGGGCTATCTGTTTGTAAAGCAGAGACTTGTACT 1005
QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyValAlaSerAsnPheTyrlleArgArg 160
DB CAAAATAATTCAGTACGATGCTGATGCTGATGCTGCAAGTAACTTCTATATCCGAGA 1065
QY 161 AlaLeuGlnIleLeuThrAsnAlaIleThrGlnCysThrLeuLeuCysProProProArgLeu 180
DB GCTCTGAAAATTTTAAACAAACGACACAGTGCATTTGTTGTCTCTCCCAAGACTA 1125
QY 181 CysThrAspAsnGlyIleMetIleAlaArgAsnGlyIleGluArgLeuArgAlaGlyLeu 200
DB TGCACGTATATGCGATTATGATTCATGATGCAATGTAATGTAAGAACTACGCTGGCTTG 1185
QY 201 GlyIleLeuHisAspIleGluGlyIleArgTyrlleProLysCysProLeuGlyValAsp 220
DB GGCATTTTATCATGACATAGAAAGCATCCGCTATGAAACCAAAATGCTCTTGGAGTAGAC 1245
QY 221 HisSerLysGluValGlyLysIleAsnIleLysValProGlnLeuLysMetGluIle 239
DB 1246 AATATCAAAAGAAAGTGGAAAGCTTCATATAAAGTACCAATTTAAAAATGAGATA 1302
RESULT 2
US-10-067-443-23

Sequence 23, Application US/10067443
Publication No. US20030082782A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEINASE HIGHLY EXPRESSED IN
TITLE OF INVENTION: SPINAL CORD, MP-1
FILE REFERENCE: D0073 NP
CURRENT APPLICATION NUMBER: US/10/067,443
CURRENT FILING DATE: 2002-02-05
PRIOR APPLICATION NUMBER: US 60/266,518
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: US 60/282,814
PRIOR FILING DATE: 2001-04-10
NUMBER OF SEQ ID NOS: 71
SOFTWARE: PatentIn version 3.0
SEQ ID NO 23
LENGTH: 1526
TYPE: DNA
ORGANISM: homo sapiens
US-10-067-443-23

Alignment Scores:

Pred. No.:	1,646-151	Length:	1526
Score:	1240.00	Matches:	239
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-10-649-273-2_COPY_176_414 (1-239) x US-10-067-443-23 (1-1526)

QY 1 LeuLeuAlaLeuValGInGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
DB CTGTTGGCATTAGTTCAGAGAGATTTCAGATTTCCTGCTTGGAAAGTCTTGGACATA 144
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuLeuIleLysIleProGlu 40
DB GACCCAGGTGACATGCTTGAAGAAGTGGCAAGAGACCTTCTTAAATAAACATCCAGAG 204
QY 41 CysSerThrMetSerGlyLysAlaIleGluHisLeuAlaLysGlnGlyAsnArgPhe 60
DB TGTCTCCACCATGAGTGGGAAAAGCCATGAAACATTTGGCCAAACAGAAATAGATT 264
QY 61 HisPheAspIleLysProProLeuHisIleAlaLysAsnCysAspPheSerPheThrGly 80
DB CATTTGACATCAAACTCCCTTCATCATGCTAAATAATGCTATTTCTTTTACTGGA 324
QY 81 LeuGlnHisValThrAspLysIleIleMetLysLysGluLysGluGluGlyIleGluLys 100
DB CTTCAACACCTTCTGATTAATAATATGAAAAAGAAAAAGAGATTGACAAAG 384
QY 101 GlnGlnIleLeuSerSerAlaIleAspIleAlaIleArgValGlnHisIleMetAlaCys 120
DB CTTCAACACCTTCTGATTAATAATATGAAAAAGAAAAAGAGATTGACAAAG 444
QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
DB GGGCAAAATCTGCTTCAGCAGCAGACATGCTGCCACAGTACGACACACATGGCATGT 444
QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyValAlaSerAsnPheTyrlleArgArg 160
DB CAAAATAATTCAGTACGATGCTGATGCTGATGCTGCAAGTAACTTCTATATCCGAGA 504
QY 161 AlaLeuGlnIleLeuThrAsnAlaIleThrGlnCysThrLeuLeuCysProProProArgLeu 180
DB GCTCTGAAAATTTTAAACAAACGACACAGTGCATTTGTTGTCTCTCCCAAGACTA 624
QY 181 CysThrAspAsnGlyIleMetIleAlaArgAsnGlyIleGluArgLeuArgAlaGlyLeu 200
DB TGCACGTATATGCGATTATGATTCATGATGCAATGTAATGTAAGAACTACGCTGGCTTG 684
QY 201 GlyIleLeuHisAspIleGluGlyIleArgTyrlleProLysCysProLeuGlyValAsp 220

Db	685	GGCAATTACATGATGACATGAGAGGCAATCCGCTAATGAACCAAAATGTCCTCTTGAGATGAC	74
Qy	221	IISeerlysglVAlGlyGlnAlaSerIlelysvaIProGlnleuIysetGluIle	239
Db	745	ATATCAAAAGAAAGTTGGAGAAAGCTTCATTAAGTACCAAAATTAATGAAGATA	801
RESULT 3			
US-10-649-273-23			
Sequence 23, Application US/10649273			
Publication No. US20040043407A1			
GENERAL INFORMATION:			
APPLICANT: Bristol-Myers Squibb Company			
TITLE OF INVENTION: POLYPEPTIDE ENCODING A NOVEL METALLOPROTEASE, MP-1			
FILE REFERENCE: D0073 CNT			
CURRENT APPLICATION NUMBER: US/10/649,273			
CURRENT FILING DATE: 2003-08-27			
PRIOR APPLICATION NUMBER: US 60/266,518			
PRIOR FILING DATE: 2001-02-05			
PRIOR APPLICATION NUMBER: US 10/067,443			
PRIOR FILING DATE: 2002-02-05			
PRIOR APPLICATION NUMBER: US 60/282,814			
PRIOR FILING DATE: 2001-04-10			
NUMBER OF SEQ. ID NOS: 71			
SOFTWARE: PatentIn version 3.2			
SEQ. ID NO. 23			
LENGTH: 1526			
TYPE: DNA			
ORGANISM: homo sapiens			
US-10-649-273-23			
Alignment Scores:			
Pred. No.:	1,64e-151	Length:	1526
Score:	1240.00	Matches:	239
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	18	Gaps:	0
US-10-649-273-2, COPY_176_414 (1-239) x US-10-649-273-23 (1-1526)			
Qy	1	LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlnGlySerLeuAspIle	20
Db	85	CTGTTGGCATTTAGTTCAAGAGGTTTCAGATTTCTGCTTCTGGAAAGCTTTGGACATA	144
Qy	21	AlaProGlyAspMetLeuAspIlysvaIAlaIArgArgLeuSerIlelyshIspTroGlu	40
Db	145	GCACAGGAGCATGCTTGACAAAGTGGCAAGAAAGCTTTCTTAATAAACATCCAGAG	204
Qy	41	CysSerThrMetSerGlyIlysvaIleGlnHisLeuAlaIysGlnGlnIyAsnArgPhe	60
Db	205	TGCTCCACCATATGATGTGGAAAGCCATAGAACATTTGGCCAAACAGGAATTAAGATTT	264
Qy	61	HisPheAspIleIysProPheLeuHisIshIAlaIysAsnCysAspPheSerPheThrGly	80
Db	265	CATTTTGACATCAAACTCCCTTGACATCAATCTTAATAATTTGATTTTCTTTTACTGGA	324
Qy	81	LeuGlnHisValThrAspIlyshIleIleMetIyLysGlnIyGlnGlnGlyIleGlnIyAs	100
Db	325	CTTCAACAGCTTATCATATAAATAATTAAGAAAAGAAAAGAGAGAGGATTTGAGAG	384
Qy	101	GlyGlnIleLeuSerSerAlaAlaAspIleAlaIleThrValGlnHisThrMetAlaCys	120
Db	385	GGGCAAAATCTGTCTTCAGCAGCAGCATTTGTGCCACAGTACGACACAAATGGCAATGT	444
Qy	121	HisLeuValIysArgThrHisArgAlaIleLeuPheCysIyGlnArgAspLeuLeuPro	140
Db	445	CATCTTGTAAGAAAGAACACATCGGCTATTTCTGTTTGTAAAGCAGAGACTTGTAACCT	504
Qy	141	GlnAsnAsnAlaValLeuValAlaSerGlyIlyValAlaSerAsnPheTyrlIeArgArg	160
Db	505	CAAAATATATGACAGTACGTGTGATCATCGTGGTGTGCAAGTAATCTTATATCCGAGA	564
Qy	161	AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrIleuLeuCysProProProArgLeu	180

```

Db      565  GCTCTGGAATTTTAAACAAACGCAACACAGTCACCTTGTGTGTCCTCTCCCAACTA 624
Oy      181  CysThrAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuAlaGlyLeu 200
Db      625  TGCACGTGTAATGCGATTATGATTCATGGAATGGATGGAAGACTACGTCGCTTG 684
Oy      201  GlyIleLeuHisAspIleGluGlyIleArgTyrGluProLysCysPheLeuGlyValAsp 220
Db      685  GGCATTTTACATGACATAGAAAGGCATCCGCTATGAACCAAAAGTCTCTTGGAGTAGAC 744
Oy      221  ILeSerLysGluValGlyGluIleSerIleLysValProGluLeuLysMetGluIle 239
Db      745  ATATCAAAAGAGAGTTGGAGAAAGCTTCATTAAGTACCAATTAATAATGGAGATTA 801

RESULT 4
US-10-651-722-23
; Sequence 23, Application US/10651722
; Publication No. US20040048302A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE, MP-1
; FILE REFERENCE: D0073 DIV
; CURRENT APPLICATION NUMBER: US/10/651,722
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 10/067,443
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: Patentin version 3.2.
; SEQ ID NO 23
; LENGTH: 1526
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-651-722-23

```

Alignment Scores:			
Pred. No.:	1,64e-151	Length:	1526
Score:	1240.00	Matches:	239
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	18	Gaps:	0
US-10-649-273-2_COPY_176_414 (1-239) x US-10-651-722-23 (1-1526)			
Qy	1	LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuLeuGlyLysSerLeuAspIle	20
Db	85	CTGTTGGCATTAGTTCAAGAGATTTCAGATTTTCTGCTTCTTGAAAGTCCTTGACATA	144
Qy	21	AlaPArgLysAspMetLeuAspLysValAlaAArgLysSerLeuLileYshSPArgLu	40
Db	145	GCACGAGGAGACATGCTTGACAAAGGCGAGAAAGACTTTCTTAAATAAACATCCAGAG	204
Qy	41	CysSerThrMetSerGlyGlyLysAlaIleGlnHisLeuAlaLysGlnGlyAsnArgPhe	60
Db	205	TGCTCCACCATATGATGCTGGAGAAAGCCATAGAACATTTGGCCCAACACAGAAATAGATTT	264
Qy	61	HisPheAspLileYsPArgProLeuHisHisAlaLysAsnCysAspPheSerPheThrGly	80
Db	265	CATTTTGATCATCAACCTCCCTTGACATCAACGTAAATAATGTATTTTCTTTATCTGGA	324
Qy	81	LeuGlnHisValThrAspLysIleIleMetLysLysGluYsgIugIugIyIleGluLys	100
Db	325	CTTCAACACGTTACGTATATAAATAAGCAAAAAGGAAAGGAGGAAAGGTATTTGAGAG	384
Qy	101	GlyGlnIleLeuSerSerAlaAlaAspLileAlaAlaThrValGlnHisSThrMetAlaCys	120
Db	385	GGGCAAACTCTCTTTCAGACAGACATTTGCTGCCACATGACAGCACAAATGGCATGT	444

QY 121 HisLeuValIysArgThrHisArgAlaIleLeuPheCysIysGlnArgAspLeuPro 140
DB 445 CATCTGTGAAAAAGAACACATCGGGCTATTCTGTTTGTGAAGCAGACAGACTTGTTACT 504
QY 141 GlnAspAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnProTyrIleArgArg 160
DB 505 CAAATAAATGACGACTGCTGTGATCTGTGTGTGTCGAGTAACCTTCATATCCGAGAGA 564
QY 161 AlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
DB 565 GCTCTGAAATTTTAACAACGACACAGACAGCTTGTGTGTCTCTCCACAGACTA 624
QY 181 CysThrAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeu 200
DB 625 TGCACTGATATATGCGCATTAATGATTGCAATGTAATGTAAGACTACAGCGCTGCTTG 684
QY 201 GlyIleLeuHisAspIleGluGlyIleArgTyrGluProIysCysProLeuGlyValAsp 220
DB 685 GGCATTTCATACATAGACAGAGCATCCGCTATGAACCAAAATGTCTCTTGAGTAGAC 744
QY 221 ILeSerIysGluValGlyGluAlaSerIleLeuValProGlnLeuIysMetGluIle 239
DB 745 ATATCAAAAGAGTGGAGAGGCTTCATATAAGTACACAAATTAATAATGAGATA 801

RESULT 5

US-10-067-443-1
; Sequence 1, Application US/10067443
; Publication No. US20030082782A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE HIGHLY EXPRESSED IN
; FILE REFERENCE: D0073 NP
; CURRENT APPLICATION NUMBER: US/10/067,443
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2197
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (231)..(1472)
US-10-067-443-1

Alignment Scores:

Pred. No.: 2,796-151 Length: 2197
Score: 1240.00 Matches: 239
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x US-10-067-443-1 (1-2197)

QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyIysSerLeuAspIle 20
DB 756 CTGTTGGCATTAATCAAGAGTTTCAGATTTCCTCTTGAAAGCTCTTGACACTA 815
QY 21 AlaProGlyAspMetLeuAspIysValAlaArgArgLeuSerLeuIleIysHisProGlu 40
DB 816 GACACGAGTGACATGCTTGACAGAGTGCGCAAGAACTTTCTTAATAAACAATCCAGAG 875
QY 41 CysSerThrMetSerGlyGlyIysValAlaIleGluHisLeuAlaIysGlnGlyAsnArgPhe 60
DB 876 TGGTCCACCATGAGTGGTGGAGAGCCATAGAACTTTGGCCAAACAGAAATAGATT 935
QY 61 HisPheAspIleIysProProLeuHisHisAlaIysAsnCysAspPheSerPheThrGly 80

DB 936 CATTTTAACATCAAACTCCCTTGCAATCATGCTAAATATGATTTTCTTTTAACAGGA 995
QY 81 LeuGlnHisValThrAspIysIleIleMetIysIysGluIysGluGlyIleGluIys 100
DB 996 CTTCACACGTTTACTGATTAATATATATGAAAAGAAAAGAGAAAGATTTAGAGAG 1055
QY 101 GlyGlnIleLeuSerSerAlaIaAspIleAlaAlaThrValGlnHisThrMetAlaCys 120
DB 1056 GGGCAATTCCTGCTTCAGAGCAGACATTCCTCCACAGTACAGACACAATGGCAGT 1115
QY 121 HisLeuValIysArgThrHisArgAlaIleLeuPheCysIysGlnArgAspLeuPro 140
DB 1116 CATCTTGAAAAGAACACATCGGGCTATTCTGTTTGTGAAGCAGAGACTTGTTACT 1175
QY 141 GlnAspAsnAlaValIleuValAlaSerGlyGlyValAlaSerAsnProTyrIleArgArg 160
DB 1176 CAAATATATGACGACTGCTGTGATCTGTGTGTGTCCAGTAACTTCATATCCGAGAGA 1235
QY 161 AlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
DB 1236 GCTCTGAAATTTTAACAACGACACAGGCACTTGTGTGTCTCTCCACAGACTA 1295
QY 181 CysThrAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeu 200
DB 1296 TGCACTGATATATGCGCATTAATGATTGCAATGTAATGTAAGACTACAGCTGCTGCTTG 1355
QY 201 GlyIleLeuHisAspIleGluGlyIleArgTyrGluProIysCysProLeuGlyValAsp 220
DB 1356 GGCATTTCATACATAGACAGAGCATCCGCTATGAACCAAAATGTCTCTTGAGTAGAC 1415
QY 221 ILeSerIysGluValGlyGluAlaSerIleLeuValProGlnLeuIysMetGluIle 239
DB 1416 ATATCAAAAGAGTGGAGAGGCTTCATATAAGTACACAAATTAATAATGAGATA 1472

RESULT 6

US-10-649-273-1
; Sequence 1, Application US/10649273
; Publication No. US20040043407A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE, MP-1
; FILE REFERENCE: D0073 CNT
; CURRENT APPLICATION NUMBER: US/10/649,273
; CURRENT FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 10/067,443
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2197
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (231)..(1472)
US-10-649-273-1

Alignment Scores:

Pred. No.: 2,796-151 Length: 2197
Score: 1240.00 Matches: 239
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x US-10-649-273-1 (1-2197)

QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyIysSerLeuAspIle 20

```

Db      756 CTGTTGGCATTAGTTAGAGAGATTTCAGATTTCCTCTTGGAAAGCTTTGGACATA 815
Qy      21 A1APROGLYASPMETLEUASPPLYEVALAARGARGLEUSERTLEUHSIPROGLU 40
Db      816 GCACCAAGGTGACATGCTTGACAGAGGTGGCAAGAACCTTTCTTAAATTAACATCCAGAG 875
Qy      41 CysSerThrMetSerGlyGlyLysAla1EGLUHSLEUALYSGINGLYASNAARGPHE 60
Db      876 TGCTCCACCATGATGCTGTGGAAAGCCATAGAACATTTGGCCAAACAGAAATAGATTT 935
Qy      61 HisPheAsp1IleYsProProLeuHisHisAlaLysAsnCyAspPheSerPheThrGly 80
Db      936 CATTTTGACATCAAACTCCCTTGACATCATGCTAAAAATTTGATTTTCTTTTACATGGA 995
Qy      81 LeuGlnHisValThrAspIleIleMetLysLysGluLysGluGluGly1IEGLUYS 100
Db      996 CTTCAACACGTTACTATATAAATAATTAATGAAGAAAGGAAAGGATTTAGAGAG 1055
Qy      101 GlyGlnIleLeuSerSerAlaAlaAspIleAla1AthrValGlnHisThrMetAlaCys 120
Db      1056 GGGCAAAATCTGCTCTTACAGACAGACATTTGCTCCACAGTACAGACACAAATGGCATGT 1115
Qy      121 HisLeuValLysArgThrHisArgAla1IleLeuPheCysLysGlnArgAspLeuPro 140
Db      1116 CATCTTGAAAAGAACACACTGGGGCTATTCTGTTTGTAGACAGAGAGACTTGTACT 1175
Qy      141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPheTy1IleArgArg 160
Db      1176 CAAATAAATGCAAGTACTGCTGTGATCGTGGTGGCAAGTAATCTTATATCCGAGA 1235
Qy      161 AlaLeuGlu1IleuThrAsnAla1ThrGlnCysThrLeuLeuCysProProArgLeu 180
Db      1236 GCTCTGAAATTTTAAACAAACGACACAGTGACTTGTGTGCTCTCCACAGACTA 1295
Qy      181 CysThrAspAsnGlyIleMetIleAla1ArgAsnGly1IEGLUARGLUEARGALAGLYLEU 200
Db      1296 TGCACGTATATGCGCATTAATGATTCATGAAATGCTATTAAGAACCTACGCTGGCTTG 1355
Qy      201 GlyIleLeuHisAsp1IleGluGlyIleArgTyrgLuprolyCysProLeuGlyValAsp 220
Db      1356 GCATATTTACATGACATAGAGGACATCCGCTATGAACCAAAATGTCTCTTGAGATGAC 1415
Qy      221 HisSerLysGluValGlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 239
Db      1416 ATATCAAAAGAAAGTTGGAAGAGCTTCATTAAGTACCAATTAATAATGAGAGATA 1472

RESULT 7
US-10-651-722-1
; Sequence 1, Application US/10651722
; Publication No. US20040048302A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OR INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE, MP-1
; FILE REFERENCE: D0073 DIV
; CURRENT APPLICATION NUMBER: US/10/651,722
; PRIOR FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 10/067,443
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2197
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (231)..(1472)
US-10-651-722-1

```

```

Alignment Scores:
Pred. No.: 2,796-151 Length: 2197
Score: 1240.00 Matches: 239
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-649-273-2_copy_176_414 (1-239) x US-10-651-722-1 (1-2197)

Qy      1 LeuLeuAlaLeuValGlnGlyVal1SerAspPheLeuLeuGlyLysSerLeuAsp1Ile 20
Db      756 CTGTTGGCATTAGTTAGAGAGATTTCAGATTTCCTCTTGGAAAGCTTTGGACATA 815
Qy      21 A1APROGLYASPMETLEUASPPLYEVALAARGARGLEUSERTLEUHSIPROGLU 40
Db      816 GCACCAAGGTGACATGCTTGACAGAGGTGGCAAGAACCTTTCTTAAATTAACATCCAGAG 875
Qy      41 CysSerThrMetSerGlyGlyLysAla1EGLUHSLEUALYSGINGLYASNAARGPHE 60
Db      876 TGCTCCACCATGATGCTGTGGAAAGCCATAGAACATTTGGCCAAACAGAAATAGATTT 935
Qy      61 HisPheAsp1IleYsProProLeuHisHisAlaLysAsnCyAspPheSerPheThrGly 80
Db      936 CATTTTGACATCAAACTCCCTTGACATCATGCTAAAAATTTGATTTTCTTTTACATGGA 995
Qy      81 LeuGlnHisValThrAspIleIleMetLysLysGluLysGluGluGly1IEGLUYS 100
Db      996 CTTCAACACGTTACTATATAAATAATTAATGAAGAAAGGAAAGGATTTAGAGAG 1055
Qy      101 GlyGlnIleLeuSerSerAlaAlaAspIleAla1AthrValGlnHisThrMetAlaCys 120
Db      1056 GGGCAAAATCTGCTCTTACAGACAGACATTTGCTCCACAGTACAGACACAAATGGCATGT 1115
Qy      121 HisLeuValLysArgThrHisArgAla1IleLeuPheCysLysGlnArgAspLeuPro 140
Db      1116 CATCTTGAAAAGAACACACTGGGGCTATTCTGTTTGTAGACAGAGAGACTTGTACT 1175
Qy      141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPheTy1IleArgArg 160
Db      1176 CAAATAAATGCAAGTACTGCTGTGATCGTGGTGGCAAGTAATCTTATATCCGAGA 1235
Qy      161 AlaLeuGlu1IleuThrAsnAla1ThrGlnCysThrLeuLeuCysProProArgLeu 180
Db      1236 GCTCTGAAATTTTAAACAAACGACACAGTGACTTGTGTGCTCTCCACAGACTA 1295
Qy      181 CysThrAspAsnGlyIleMetIleAla1ArgAsnGly1IEGLUARGLUEARGALAGLYLEU 200
Db      1296 TGCACGTATATGCGCATTAATGATTCATGAAATGCTATTAAGAACCTACGCTGGCTTG 1355
Qy      201 GlyIleLeuHisAsp1IleGluGlyIleArgTyrgLuprolyCysProLeuGlyValAsp 220
Db      1356 GCATATTTACATGACATAGAGGACATCCGCTATGAACCAAAATGTCTCTTGAGATGAC 1415
Qy      221 HisSerLysGluValGlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 239
Db      1416 ATATCAAAAGAAAGTTGGAAGAGCTTCATTAAGTACCAATTAATAATGAGAGATA 1472

RESULT 8
US-10-480-988-36
; Sequence 36, Application US/10480988
; Publication No. US20050069877A1
; GENERAL INFORMATION:
; APPLICANT: SANDHI, Ameena R.; KABIR, Amy B.;
; APPLICANT: TRAN, Bao; DUGGAN, Brendan M.;
; APPLICANT: WARREN, Bridget A.; ISON, Craig H.;
; APPLICANT: HONCHRIE, Cynthia D.; MEYER, Daniel P.;
; APPLICANT: LU, Dyrung Aina M.; LEE, Ernestine A.;
; APPLICANT: YUE, Henry; FORSYTHE, Ian J.;
; APPLICANT: BARROSO, Ines; RAMKUMAR, Jayalaxmi;
; APPLICANT: GRIFPIN, Jennifer A.; LI, Joana X.;

```

APPLICANT: YANG, Junning; THANGAVELU, Kavitha;
APPLICANT: GIETZEN, Kimberly J.; DING, Li;
APPLICANT: BAUGHN, Mariah R.; BOROWSKY, Mark L.;
APPLICANT: YAO, Monique G.; CHAMLA, Narinder K.;
APPLICANT: MASON, Patricia M.; GURURAJAN, Rajagopal;
APPLICANT: LEE, Sally; BECHA, Shanya D.;
APPLICANT: LEE, Soo Yeun; TRAN, Uyen K.;
APPLICANT: ELIOT, Vicki S.; LEO, Wen;
APPLICANT: SPRAGUE, William W.; TANG, Y. Tom;
APPLICANT: LU, Yan; ZEBARJADIAN, Yeganeh
TITLE OF INVENTION: PROTEIN MODIFICATION AND MAINTENANCE MOLECULES
FILE REFERENCE: PF-1040 USN
CURRENT APPLICATION NUMBER: US/10/480, 988
CURRENT FILING DATE: 2003-12-16
PRIOR APPLICATION NUMBER: PCT/US02/19360
PRIOR FILING DATE: 2002-06-18
PRIOR APPLICATION NUMBER: US 60/300,508
PRIOR FILING DATE: 2001-06-22
PRIOR APPLICATION NUMBER: US 60/303,445
PRIOR FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: US 60/305,405
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/311,442
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 60/314,821
PRIOR FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/315,992
PRIOR FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: US 60/378,205
PRIOR FILING DATE: 2002-05-03
NUMBER OF SEQ ID NOS: 56
SOFTWARE: PERL Program
SEQ ID NO 36
LENGTH: 2572
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 7632424CB1
US-10-480-988-36

Alignment Scores:
Pred. No.: 3,52e-151 Length: 2572
Score: 1240.00 Matches: 239
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x US-10-480-988-36 (1-2572)

QY 1 LeuleuAlaleuValGInglYValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
DB 669 CTTGTTGGCATTGATTCAAGAGATTTCAGATTTCTGCTTTCGGAAGCTTTGGACATA 728
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 729 GCACCAAGTGACATGCTTGAACAAGTGCGCAAGAGCTTTCTTAATAAACAATCCAGAG 788
QY 41 CysSerThMetSerGlyGlyLysAlaIleGluHisLeuAlaLysGInglYAsnArgPhe 60
DB 789 TGCCTCCACCATGAGTGTGGGAAAGCCATAGAACATTTGGCCCAACAGAAATAGATTT 848
QY 61 HisPheAspIleLysProLeuHisHisAlaLysAsnCysAspPheSerPheThrGly 80
DB 849 CATTTCGACATCAACTCCCTTCGATCATGCTAATAAATGTGATTTTCTTTTACGGA 908
QY 81 LeuGlnHisValThrAspLysIleIleMetLysLysGlyLysGInglYIleGluLys 100
DB 909 CTTCAACACAGTTAATCATATAAATAAATAAAGAAAGAAAGAGATATTGAGAAC 968
QY 101 GlyGlnIleLeuSerSerAlaIleAspIleAlaIleThrValGlnHisThrMetAlaCys 120

DB 969 GGGCAAACTCTGCTTCAGCAGCAGACATTGCTCCACAGTACAGCACACAATGCGATGT 1028
QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
DB 1029 CATTCTGAAAAAGAACACATCGGGCTATTCTGTTTGTAGACAGAGACTTTGACT 1088
QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAspPheTyrlleArgArg 160
DB 1089 CAAATATATCAGTACAGTGTGATCGTGTGCTGCCAAGTAACTTATATATCCGACGA 1148
QY 161 AlaleuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
DB 1149 GCTCTGAAATTTTAACAAACGCAACAGCAGCACTTTGTGTCTCTCCACGACTA 1208
QY 181 CysThrAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuArgAlaGlyLeu 200
DB 1209 TGCATCTATATGACATTAATGATTCATGAAATGTATTTGAAACATACGCTGCGCTTG 1268
QY 201 GlyIleLeuHisAspIleGlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 239
DB 1269 GGCATTTCATGACATGACAGGCGATCGCTTAACCAAAATGTCTTGAGTAGAC 1328
QY 221 HisSerLysGlyValGlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 239
DB 1329 ATATCAAAAGAGTTGAGAGAGCTTCATATAAAGTACCAATTAATAATGAGATA 1385

RESULT 9

US-10-067-443-21
Sequence 21, Application US/10067443
Publication No. US20030082782A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE HIGHLY EXPRESSED IN
FILE REFERENCE: D0073 NP
CURRENT APPLICATION NUMBER: US/10/067,443
CURRENT FILING DATE: 2002-02-05
PRIOR APPLICATION NUMBER: US 60/266,518
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: US 60/282,814
PRIOR FILING DATE: 2001-04-10
NUMBER OF SEQ ID NOS: 71
SOFTWARE: PatentIn version 3.0
SEQ ID NO 21
LENGTH: 1387
TYPE: DNA
ORGANISM: homo sapiens
US-10-067-443-21

Alignment Scores:
Pred. No.: 4,21e-148 Length: 1387
Score: 1213.50 Matches: 238
Percent Similarity: 90.15% Conservative: 0
Best Local Similarity: 97.86% Mismatches: 1
Query Match: 97.86% Indels: 25
DB: 14 Gaps: 1

US-10-649-273-2_COPY_176_414 (1-239) x US-10-067-443-21 (1-1387)

QY 1 LeuleuAlaleuValGInglYValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
DB 549 CTTGTTGGCATTGATTCAAGAGATTTCAGATTTCTGCTTTCGGAAGCTTTGGACATA 608
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 609 GCACCAAGTGACATGCTTGAACAAGTGCGCAAGAGCTTTCTTAATAAACAATCCAGAG 668
QY 41 CysSerThMetSerGlyGlyLysAlaIleGluHisLeuAlaLysGInglYAsnArgPhe 60
DB 669 TGCCTCCACCATGAGTGTGGGAAAGCCATAGAACATTTGGCCCAACAGAAATAAGATTT 728
QY 61 HisPheAspIleLysProLeuHisHisAlaLysAsnCysAspPheSerPheThrGly 80

Db 729 CATTGGACATCAACCTCCCTGCATCATGCTAAAAATTGTCTTTCTTTACTGGA 788
Qy 81 LeuGlnHisValThrAspLysIleIleMetLysLysGluLysGluGlyIle----- 98
Db 789 CTTCAACACGTTACTGATTAATATATGAAAAAGAGAGAGATATTTTCTA 848
Qy 98 ----- 98
Db 849 ATTAGTAAAGTTGAACAGATAAATATTCCTGATTTGCTTAAAAATAGCTCATTTTC 908
Qy 99 -----GluLysGluGlnIleLeuSerSerIleIleAspIleIleIleValGln 115
Db 909 TGCAGGTATGAGAGGGCAAAATCCTGTCTTCAGCAGCAGCATTTGCGCCAGTACAG 968
Qy 116 HisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGln 135
Db 969 CACACATGCGATGCTACTCTTGGAAGAAAGAACATCGGGCTATCTGTTTGTAGCAG 1028
Qy 136 ArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsn 155
Db 1029 AGAGACTGTGTACTCTCAAAATTAATGACGTACTGTTGCATCTGGTGGTGGCAAGTAC 1088
Qy 156 PheTyrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCys 175
Db 1089 TTCTATATCCGACAGCTCTGGAATTTTAAACCAACCAACAGTGCATCTTGTGTGT 1148
Qy 176 ProProProArgLeuCysThrAspAsnGlyIleMetIleIleIleIleIleIleIleIle 195
Db 1149 CTTCTCCCAAGCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1208
Qy 196 LeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGluProLysCys 215
Db 1209 CTACGCTGCTGCTGGCATTTTACATGACATGACATGACATGACATGACATGACATGAC 1268
Qy 216 ProLeuGlyValAspIleSerLysGluValGlyLysAspIleLysValProGlnLeu 235
Db 1269 CCTCTTGAGATGACATATCAAAAGAGAGTGGAGAGCTTCATTAAGTACCAATTA 1328
Qy 236 LysMetGluIle 239
Db 1329 AAAATGAGATA 1340
RESULT 10
US-10-649-273-21
; Sequence 21, Application US/10649273
; Publication No. US20040043407A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE, MP-1
; FILE REFERENCE: D0073 CMT
; CURRENT APPLICATION NUMBER: US/10/649, 273
; CURRENT FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US 60/266, 518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 10/067, 443
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/282, 814
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 1387
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-649-273-21
Alignment Scores:
Pred. No.: 4,21e-148 Length: 1387
Score: 1213.50 Matches: 238
Percent Similarity: 50.15% Conservative: 0
Best Local Similarity: 90.15% Mismatches: 1
Query Match: 97.86% Indels: 25
DB: 18 Gaps: 1

US-10-649-273-2_copy_176_414 (1-239) x US-10-649-273-21 (1-1387)
Qy 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
Db 549 CTGTGGCATAGTATGTTCAAGAGAGTTTCAGATTTTCCTCTTGGAAAGCTTTGACATA 608
Qy 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
Db 609 GCACCAAGGTGACATGCTTGCACAGGTGCAAGAGACATTTCTTAAATAAACATCCAGAG 668
Qy 41 CysSerThrMetSerGlyGlyValAlaIleGlnHisLeuAlaLysGlnGlyValAsnArgPhe 60
Db 669 TGCCTCACCATGAGGTGGGAAAGCCATYAGACATTTGGCCAAACAGAAATAGATT 728
Qy 61 HisPheAspIleLysProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGly 80
Db 729 CATTGTGACATCAACCTCCCTTGATCATGCTAAATTTGTGATTTTCTTTTACTGGA 788
Qy 81 LeuGlnHisValThrAspLysIleIleMetLysLysGluLysGluGlyIle----- 98
Db 789 CTTCAACACGTTACTGATTAATATATGAAAAAGAGAGAGATATTTCTA 848
Qy 98 ----- 98
Db 849 ATTAGTAAAGTTGAACAGATAAATATTCCTGATTTGCTTAAAAATAGCTGCTCATTTTC 908
Qy 99 -----GluLysGluGlnIleLeuSerSerIleIleAspIleIleIleValGln 115
Db 909 TGCAGGTATGAGAGGGCAAAATCCTGTCTTCAGCAGCAGCATTTGCGCCAGTACAG 968
Qy 116 HisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGln 135
Db 969 CACACATGCGATGCTACTCTTGGAAGAAAGAACATCGGGCTATCTGTTTGTAGCAG 1028
Qy 136 ArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsn 155
Db 1029 AGAGACTGTTAACCTCAAAATTAATGACGTACTGTTGCATCTGGTGGTGGCAAGTAC 1088
Qy 156 PheTyrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCys 175
Db 1089 TTCTATATCCGACAGCTCTGGAATTTTAAACCAACCAACAGTGCATCTTGTGTGT 1148
Qy 176 ProProProArgLeuCysThrAspAsnGlyIleMetIleIleIleIleIleIleIleIle 195
Db 1149 CTTCTCCCAAGCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1208
Qy 196 LeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGluProLysCys 215
Db 1209 CTACGCTGCTGCTGGCATTTTACATGACATGACATGACATGACATGACATGACATGAC 1268
Qy 216 ProLeuGlyValAspIleSerLysGluValGlyLysAspIleLysValProGlnLeu 235
Db 1269 CCTCTTGAGATGACATATCAAAAGAGAGTGGAGAGCTTCATTAAGTACCAATTA 1328
Qy 236 LysMetGluIle 239
Db 1329 AAAATGAGATA 1340
RESULT 11
US-10-651-722-21
; Sequence 21, Application US/10651722
; Publication No. US20040048302A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE, MP-1
; FILE REFERENCE: D0073 DIV
; CURRENT APPLICATION NUMBER: US/10/651, 722
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US 60/266, 518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 10/067, 443
; PRIOR FILING DATE: 2002-02-05

DB 1089 TCTATATCCGACGCTCTGAAATTTTAAACCAACACAGTCGACTTGTGTGT 1148
QY 176 PROPROARGLEUCYSTRASPANGIYILEMETLEALATPAANGIYILEGUARG 195
DB 1149 CCCCCCAAGATATGCACTGATTAATGCAATTAATGATTAATGATTAATGATTA 1268
QY 196 LEUARGALAGIYLEUGIYILEUHIASAPILEGIUGIYILEARGIYRGUPROLYCY 215
DB 1209 CTAAGTGTGGCTTGGGCACTTTTACATGACATAGAGGCAATCCGCTATGAACCAAAATGT 1268
QY 216 PROLEUGIYVALASPILESERLYGIVUALIGIYUUALASERILEYVALPROGILEU 235
DB 1269 CCTCTTGAGATGACATATCAAAAAAGAGTTGAGAAAGCTTCCATTAACCAAAATTA 1328
QY 236 LYSMETGLUILE 239
DB 1329 AAAATGAGATA 1340
RESULT 13
US-10-012-140-6
; Sequence 6, Application US/10012140
; Publication No. US2003009017A1
; GENERAL INFORMATION:
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
; TITLE OF INVENTION: 14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 381552004900
; CURRENT APPLICATION NUMBER: US/10/012,140
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/246,768
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,772
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,185
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-140-6
Alignment Scores:
Pred. No.: 8,53e-147 Length: 1245
Score: 1203.00 Matches: 232
Percent Similarity: 98.33% Conservative: 3
Best Local Similarity: 97.07% Mismatches: 4
Query Match: 97.02% Indels: 0
DB: 14 Gaps: 0
US-10-649-273-2_COPY_176_414 (1-239) x US-10-012-140-6 (1-1245)
QY 1 LeuLeuAlaLeuValGInGlyValSerAspHeLeuLeuLeuGlyLeuSerLeuAspIle 20
DB 526 CTGTGGCATTAAGTTCAGAGGATTTCAGATTTCGTCTTGGAAAGTCTTGGACATA 585
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 586 GCACCAAGTACATGCTTGAACAGGTGCAAGAGACCTTCTTAATAAATCAATCCAGAG 645
QY 41 CysSerThrMetSerGlyGlyValAlaIleGluHisLeuAlaLysGInGlyAsnArgPhe 60
DB 646 TGTCTCCACCAAGTGGTGGAGAAAGCCATAGAACCTTGGCAACCAAGAAATAGATT 705
QY HisPheAspIleLysProProLeuHisAlaLysAsnCyAspPheSerPheThrGly 80
DB 706 CATTTTACATCAAACTCCCTTGCATCATCTAAATTTGATTTTCTTTTACATGGA 765
QY 81 LeuGlnHisValThrAspLysIleIleMetLysGlyLysGluGluGlyIleGluLys 100

DB 766 CTTCAACCGCTTACTGATTAATAATGAAGAAACGAGAAACAGAGGATTTGAGAG 825
QY 101 GIVGInIleLeuSerSerAlaAlaAspIleAlaIleValGlnHisThrMetAlaCy 120
DB 826 GGGCAATTCCTGCTTTCAGCAGCAGACATTCCTGCAAGATACAGCAATGSCATGT 885
QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
DB 886 CATCTTGAAAGAAACACATCGGGCTATCTGTTTGTAAAGCAGAGACTGTTTACT 945
QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyValAlaIleSerAspPheThrIleArgArg 160
DB 946 CAATAATATGACATGCTGTTGACATCTGCTGTGCGAAGTAACCTTATATTCGACGA 1005
QY 161 AlaLeuGluIleLeuThrAsnAlaThrGlnCySerThrLeuLeuCyProProArgLeu 180
DB 1006 GCTCTGAAATTTTAAACCAACGACACACAGTCGACTTGTGTCTCTCCACGACTA 1065
QY 181 CysThrAspAngIyIleMetIleAlaIlePaAngIyIleGuArgLeuArgAlaGlyLeu 200
DB 1066 TGCACTGATATGCACTTATGATTCATGATGATGATGATGATGATGATGATGATG 1125
QY 201 GIVGInIleLeuHisAspIleGluGlyIleArgGlyGluProLysCyProLeuGlyValAsp 220
DB 1126 GGCATTTTACATGACATAGAGGATCCGCTATGAACCAAAATGCTCTTGGAGTAC 1185
QY 221 ILeSerLysGluValGlyGluAlaSerIleLysValProGlnLeuLysMetGluIle 239
DB 1186 ATATCAAAAGAGATTGAGAGAGCTTCCATTAAGTACCAATTAATAATGAGATA 1242
RESULT 14
US-10-012-140-4
; Sequence 4, Application US/10012140
; Publication No. US2003009017A1
; GENERAL INFORMATION:
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
; TITLE OF INVENTION: 14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 381552004900
; CURRENT APPLICATION NUMBER: US/10/012,140
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/246,768
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,772
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,185
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1820
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (146)...(1390)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1820)
; OTHER INFORMATION: n = A,T,C or G
US-10-012-140-4
Alignment Scores:
Pred. No.: 1,49e-146 Length: 1820
Score: 1203.00 Matches: 232
Percent Similarity: 98.33% Conservative: 3
Best Local Similarity: 97.07% Mismatches: 4
Query Match: 97.02% Indels: 0
DB: 14 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x US-10-012-140-4 (1-1820)

QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
Db 671 CTGTTGGCATTATCTCAAGAGGTTTCAGATTTCTGCTTTCGAAAGCTTTGAGATA 730

QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
Db 731 GCAACGAGTGACATGCTTGCAAGAGTGCGCAAGAACTTTCTTAATAAACAATCCAGAG 790

QY 41 CysSerThrMetSerGlyGlyLysAlaIleGlnHisLeuAlaLysGlnGlyAspArgPhe 60
Db 791 TGTCTCCACCATGAGTGTGGGAAAGCCATAGAACATTGTGCGCAACAGAAATAGATTT 850

QY 61 HisPheAspIleLysProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGly 80
Db 851 CATTTCGACATCAACCTCCCTTCGATCAGCTAATAAATGTGATTTTCTTTTACTGGA 910

QY 81 LeuGlnHisValThrAspLysIleIleMetLysGlyLysGlnGlyGluGlyIleGluLys 100
Db 911 CTTCAACAGCTTATCTGATATAAATAATGAAACAGAAACAGAAAGAGATTTGAGAG 970

QY 101 GlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCys 120
Db 971 GGGCAATCTCTGCTTCAGCAGACGATTGCTGCCACATGACACACAAATGCGCATGT 1030

QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
Db 1031 CATCTTGTAAGAAAGAACACATCGGGCTATCTCTGTTTGTAAAGCAGAGACTTGTACT 1090

QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPheTyrIleArgArg 160
Db 1091 CAAATAATAGCAGTACTGTTGATCATCGTGTGTGCGCAAGTAATCTTATATCCGAGA 1150

QY 161 AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLeu 180
Db 1151 GCTCTGGAATTTTAAACAACGACACAGTGCATTTGTTGTCTCTCCACAGACTA 1210

QY 181 CysThrAspAsnGlyIleMetIleAlaThrAsnGlyIleGlnArgLeuArgAlaGlyLeu 200
Db 1211 TGCATCGATTAATGCGCATTAATGATTCATGGAATGTAATGAAGACTACAGCTGCGCT 1270

QY 201 GlyIleLeuHisAspIleGlnGlyIleArgTyrGlnProLysCysProLeuGlyValAsp 220
Db 1271 GGCATTTTACATGACATAGAGCACTCCGCTATGAGCAACAAATCTCTCTTGAGTGA 1330

QY 221 HisSerLysGluValGlyGlnAlaSerIleLysValProGlnLeuLysMetGluIle 239
Db 1331 ATATCAAAAGAGGTGGAAGAGCTTCATAAAGTACCAATTAATAATGAGAGATA 1387

RESULT 15
US-10-094-749-400
; Sequence 400, Application US/10094749
; Publication No: US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YORI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI

APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 400
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-400

Alignment Scores:
Pred. No.: 1,42e-127 Length: 2208
Score: 1059.00 Matches: 211
Percent Similarity: 88.28% Conservative: 0
Best Local Similarity: 88.28% Mismatches: 4
Query Match: 85.40% Indels: 24
DB: 17 Gaps: 1

US-10-649-273-2_COPY_176_414 (1-239) x US-10-094-749-400 (1-2208)

QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
Db 869 CTGTTGGCATTATCTCAAGAGGTTTCAGATTTCTGCTTTCGAAAGCTTTGAGATA 928

QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
Db 929 GCAACAGTGACATGCTTGAACAAGTGCGCAAGAAAGACTTCTTAATAAATCATCAGAG 988

QY 41 CysSerThrMetSerGlyGlyLysAlaIleGlnHisLeuAlaLysGlnGlyAspArgPhe 60
Db 989 TGTCTCCACCATGAGTGTGGGAAAGCCATAGAACATTGTGCGCAACAGAAATAGATTT 1048

QY 61 HisPheAspIleLysProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGly 80
Db 1049 CATTTCGACATCAACCTCCCTTCGATCAGCTAATAAATGTGATTTTCTTTTACTGGA 1108

QY 81 LeuGlnHisValThrAspLysIleIleMetLysGlyLysGlnGlyGluGlyIleGluLys 100
Db 1109 CTTCAACAGCTTATCTGATATAAATAATGAAAGAAAGAGAGCTTATGAGAG 1168

QY 101 GlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCys 120
Db 1169 GGGCAATCTCTGCTTCAGCAGACGACATTTCTCCACAGTACAGACACAAATGCGCATGT 1228

QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
Db 1229 CATCTTGTAAGAAACACATCGGGCTATCTGTTTGTAAAGCAGAGACTTGTACT 1288

QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPheTyrIleArgArg 160
Db 1289 CAAATAATGCAAGTCTGTTGCAATCTGCTGTGTGCGCAAGTAATCTTATATCCGACGA 1348

QY 161 AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLeu 180
Db 1349 GCTCTGGAATTTTAAACAACAGCAAGGCACTTGTGTGTGCTCTCCACAGACTA 1408

QY 181 CysThrAspAsnGlyIleMetIleAlaThrAsnGlyIleGlnArgLeuArgAlaGlyLeu 200
Db 1409 TGCATCGATTAATGCAATTAATGATTTGCA----- 1435

QY 201 GlyIleLeuHisAspIleGlnGlyIleArgTyrGlnProLysCysProLeuGlyValAsp 220
Db 1436 -----TCATGCTCTTGTGAGTAC 1456

QY 221 HisSerLysGluValGlyGlnAlaSerIleLysValProGlnLeuLysMetGluIle 239
|||

Db 1457 ATATCAAAAGAGTTGGAGAGCTTCATTAAGTACCACATTTAAAAATGGAGATA 1513

Search completed: June 17, 2005, 08:03:30
Job time : 1634.64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 16, 2005, 18:30:49 ; Search time 205.65 Seconds
(without alignments)
3294.036 Million cell updates/sec

Title: US-10-649-273-2

Perfect score: 2125
Sequence: 1 MILITKAGVFFKSRKRVY.....DISKEVGEASIKVQLMEI 414

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=x1p
-Q=/cgn2_1/USPTO.spool_p/US10649273/runat_15062005_111417_6057/app_query.fasta_1.1429
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10649273.@CGN_1.1.177 @runat_15062005_111417_6057 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfile1.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2125	100.0	2197	US-10-067-443-1	Sequence 1, Appli
2	2090.5	98.4	1387	US-10-067-443-21	Sequence 21, Appl
3	1747	82.2	1416	US-09-774-528-177	Sequence 177, App
4	1385	65.2	1526	US-10-067-443-23	Sequence 23, Appl
5	1186.5	55.8	14364	US-10-067-443-20	Sequence 20, Appl
6	502	23.6	94750	US-09-536-002-38	Sequence 38, Appl
7	501	23.6	1053	US-09-540-236-806	Sequence 806, App
8	499.5	23.5	1206	US-09-252-991A-801	Sequence 801, App
9	452.5	23.2	1059	US-09-252-991A-884	Sequence 884, App
10	482.5	22.7	1074	US-09-543-681A-2341	Sequence 2341, Ap
11	482	22.7	996	US-09-902-540-6612	Sequence 6612, Ap
12	482	22.7	2582	US-09-902-540-503	Sequence 503, App

13	479	22.5	1315	1	US-08-087-797-1	Sequence 1, Appli
14	475.5	22.4	1830121	4	US-09-557-884-1	Sequence 1, Appli
15	475.5	22.4	1830121	4	US-09-643-990A-1	Sequence 1, Appli
16	474.5	22.3	1032	4	US-09-489-039A-2050	Sequence 2050, Ap
17	472	22.2	1092	4	US-09-107-532A-2955	Sequence 2955, Ap
18	461	21.7	1008	3	US-08-987-121A-5	Sequence 5, Appli
19	461	21.7	1011	3	US-08-987-121A-3	Sequence 3, Appli
20	457	21.5	1006	3	US-08-961-083-51	Sequence 51, Appl
21	457	21.5	1006	4	US-09-536-784-51	Sequence 51, Appl
22	457	21.5	10974	3	US-08-961-527-214	Sequence 214, App
23	455	21.4	1011	3	US-09-066-512-1	Sequence 1, Appli
24	454	21.4	1101	4	US-09-134-000C-1551	Sequence 1551, Ap
25	453	21.3	1011	4	US-09-583-110-2196	Sequence 2196, Ap
26	448	21.1	1011	4	US-09-107-433-1618	Sequence 1618, Ap
27	443	20.8	1107	3	US-09-134-001C-1072	Sequence 1072, Ap
28	433	20.4	1026	3	US-09-149-624-1	Sequence 1, Appli
29	424	20.0	15249	4	US-08-956-171E-102	Sequence 102, App
30	424	20.0	15249	4	US-08-781-986A-102	Sequence 102, App
31	406.5	19.1	640681	4	US-09-790-988-1	Sequence 1, Appli
32	406.5	19.1	123025	4	US-09-198-452A-1	Sequence 1, Appli
33	406.5	19.1	1230230	4	US-09-438-185A-1	Sequence 1, Appli
34	404.5	19.0	1155	4	US-09-602-777A-147	Sequence 147, App
35	404	19.0	3064	3	US-09-221-017B-794	Sequence 794, App
36	379.5	17.9	4403765	3	US-09-103-840A-2	Sequence 2, Appli
37	379.5	17.9	4411529	3	US-09-103-840A-1	Sequence 1, Appli
38	369.5	17.4	845	4	US-09-328-352-483	Sequence 483, App
39	362	17.0	876	4	US-09-724-623-19	Sequence 19, Appl
40	357	16.8	580073	4	US-08-545-528D-1	Sequence 1, Appli
41	333.5	15.7	42325	4	US-08-311-721A-131	Sequence 131, App
42	327.5	15.4	3993	4	US-09-710-279-3985	Sequence 3985, App
43	319	15.0	822	4	US-09-710-279-727	Sequence 727, App
44	317	14.9	36941	4	US-08-311-731A-130	Sequence 130, App
45	313	14.7	1664976	4	US-08-916-421B-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-10-067-443-1
; Sequence 1, Application US/10067443
; Patent No. 6642041

GENERAL INFORMATION:

APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE HIGHLY EXPRESSED IN
TITLE OF INVENTION: SPINAL CORD, ME-1

FILE REFERENCE: D0073 NP
CURRENT APPLICATION NUMBER: US/10/067,443

CURRENT FILING DATE: 2002-02-05
PRIOR APPLICATION NUMBER: US 60/266,518

PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: US 60/282,814

PRIOR FILING DATE: 2001-04-10
NUMBER OF SEQ ID NOS: 71

SOFTWARE: PatentIn version 3.0
SEQ ID NO 1

LENGTH: 2197

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (231)..(1472)

US-10-067-443-1

Alignment Scores:

Pred. No.: 5,72e-261 Length: 2197

Score: 2125.00 Matches: 414

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 4 Gaps: 0

US-10-649-273-2 (1-414) x US-10-067-443-1 (1-2197)

Oy	1	MetLeuLiLeuThrThyStrhLaGlyValAlPhePheLysProSerLysArgLysValLyr	20
Db	231	ATGCTAACTTTCACCTAAGACCTGACGAGATTCTTTTAAACCATCAAAAAGAAAGTTTAT	290
Oy	21	GIuPheLysArgSerPheAsnPheHisProGlyThrLeuPheLeuHisLysIleValLeu	40
Db	291	GAATTTTAAAGAAAGTTTATATTTTCACTCTGGAAACATATTCTTCATATAAATAGTATTG	350
Oy	41	GIyIleGIuThrSerCyAspAspThrAlaAlaValAlaAspGIuThrGIyAsnVal	60
Db	351	GGAAATGGAACATAGTTGTGATGATATACAGACGCTGCTGTGTGATGATAAATCGAATATGTG	410
Oy	61	LeuGIyGIuAlaIleHisSerGlnThrGIuValHisLysLysThrGIyGIyIleValPro	80
Db	411	TTGGAGAGAACATATCATTTCCCAACTGAAGTTCAATTAAAAACAGTGGAGATTGTCTCT	470
Oy	81	ProAlaAlaGlnGlnLeuHisArgGIuAsnIleGlnArgIleValGlnGIuAlaLeuSer	100
Db	471	CCAGCAGGCTCAACAGCTTCACAGAGAAATATTCAACGAATAGTACAGAGCTCTTTCT	530
Oy	101	AlaSerGIyValSerProSerArgPheLysSerAlaIleAlaThrThrIleLysProGIyLeu	120
Db	531	GCCAGTGAAGTCTCTCCACAGTACCCTCAGCAATTGCAACTACCAATAAACAGAGACTT	590
Oy	121	AlaLeuSerLeuGIyValGIyLeuSerPheSerLeuGlnLeuValGIyGlnLeuLysLys	140
Db	591	GCTTTAAGCTGTGGAGTGGCTTATATCTATTAGCTTACACTGGTAGACAGTTAAAAAG	650
Oy	141	ProPheIleProIleHisIsmetGluAlaHisAlaLeuThrIleArgLeuThrAsnLys	160
Db	651	CCATTCAATCCCATTCATCATATGAGAGCTCATGCTACCTATATAGCTTGACCAATAAA	710
Oy	161	ValGIuPheProPheLeuValLeuLeuIleSerGIyHisCysLeuLeuAlaLeuVal	180
Db	711	GTAGAAATTCCTTTTATAGTCTTTTGAATTTCTGAGAGTCACTGCTGTTGGCATTAGTT	770
Oy	181	GlnGIyValSerAspPheLeuLeuLeuGIyLysSerLeuAspIleAlaProGIyAspMet	200
Db	771	CAAGGAGTTTCAGATTTCCTGTCTTGGAAAGTCTTTGGACATACACACAGGTGACATG	830
Oy	201	LeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGIuLysSerThrMetSer	220
Db	831	CTTGCAAGGTGGCAGAAAGACTTCTTTAATTAACATCCAGAGTCTCTCACCAATGAGT	890
Oy	221	GIyGIyLysAlaIleGIuHisLeuAlaLysGlnGIyAsnArgPheHisPheAspIleLys	240
Db	891	GGTGGAAAGCCATGAAACATTGGCCAAACAAGAAATAGATTCAATTTTGACATCAAA	950
Oy	241	ProProLeuHisHisAlaLysAsnCyAspPheSerPheThrGIyLeuGlnHisValThr	260
Db	951	CCTCCCTTGCAATCAGCTAAAAATGTGATTTTTCTTTTATCGGACCTTCAACACGGTTACT	1010
Oy	261	AspLysIleIleMetLysLysGIuLysGIuGlnGIyIleGIuLysGIyGlnIleLeuSer	280
Db	1011	GATTAATATATATGAAAAAGAAAAAGAAAGAGATTTGAGAGGGGCAATCTGTCT	1070
Oy	281	SerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCysHisLeuValLysArg	300
Db	1071	TCAGCAGAGACATTGCTGCCACAAGTACAGCACACAAATGGCATGTCTTGGAAAAAG	1130
Oy	301	ThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaVal	320
Db	1131	ACACATCGGGCTATTCTGTTTGTATGAACGAGAGACTTGTATCTCAAAAATATATGACATA	1190
Oy	321	LeuValAlaSerGIyGIyValAlaSerAsnPheLysIleArgArgAlaLeuGlnIleLeu	340
Db	1191	CTGTGTGACATCGGTGTGTGCAAGTAACTTTATATCCGACAGCTCTGGAAATTTTA	1250
Oy	341	ThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeuCysThrAspAsnGIy	360
Db	1251	ACAAACGCAACACAGTGCACCTTTGTGTCTCTCTCCACAGACTATGCACTGATTAATGGC	1310
Oy	361	IleMetIleAlaTrpAsnGIyIleGluArgLeuArgAlaGIyLeuGlnIleLeuHisAsp	380

Db	1311	ATTAGATTGCATGCATGATGATGATTAAGACATACGTCGCTGGCATTTATCATAC	1370
Qy	381	leugllylleargtyrsluprolysCyaproleuglyvalaspl eserlygluval	400
Db	1371	ATGAGAGCATCGCGTATGAACCAAAATGCTCTTGAGTAGACATATCAAAAGAGTT	1430
Qy	401	gllyualaser elysvalProgl leu yemerc ule	414
Db	1431	GCAGAGCTTCCATATAAAGTACCAATTAAAAATGAGATA	1472
RESULT 2			
	US-10-067-443-21		
	Sequence 21, Application US/10067443		
	Patent No. 6642041		
	GENERAL INFORMATION:		
	APPLICANT: Bristol-Myers Squibb Company		
	TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE HIGHLY EXPRESSED IN		
	TITLE OF INVENTION: SPINAL CORD, MP-1		
	FILE REFERENCE: D0073 NP		
	CURRENT APPLICATION NUMBER: US/10/067,443		
	CURRENT FILING DATE: 2002-02-05		
	PRIOR APPLICATION NUMBER: US 60/266,518		
	PRIOR FILING DATE: 2001-02-05		
	PRIOR APPLICATION NUMBER: US 60/282,814		
	PRIOR FILING DATE: 2001-04-10		
	NUMBER OF SEQ ID NOS: 71		
	SOFTWARE: PatentIn version 3.0		
	SEQ ID NO 21		
	LENGTH: 1387		
	TYPE: DNA		
	ORGANISM: homo sapiens		
	US-10-067-443-21		
Alignment Scores:			
	Pred. No.:	6.62e-257	Length: 1387
	Score:	2090.50	Matches: 412
	Percent Similarity:	93.85%	Conservative: 0
	Best Local Similarity:	93.85%	Mismatches: 2
	Query Match:	98.38%	Indels: 25
	DB:	4	Gaps: 1
US-10-649-273-2 (1-414) x US-10-067-443-21 (1-1387)			
Qy	1	Medleu leuThr ySthr a a g y val Phe phe yS ProSer y sar y y val Tyr	20
Db	24	ATGCTAATCTTACATCACTGACGAGAGGTTTTTTTAAACATCAAAAGAAAGATTAT	83
Qy	21	sluphe leu ar Ser Phe asn Phe n s Pro G Y Thr leu Phe leu n s y s le val leu	40
Db	84	GAAATTTTAAAGAAAGTTTAAATTTTCATCTCGAAACCTAATTTCTCTAAATAATGATATG	143
Qy	41	gl y le g u Thr Ser Cy asp Asp Thr a a a a val a a Phe g u Thr g y Asn val	60
Db	144	GGAAATTCGAACACTGTTGTATGATACACGACGCTGCTGTGGATGAAACTGGAATGTG	203
Qy	61	leu g y gl u a a l eh s Ser g n Thr g u val n s leu y S Thr g y le val Pro	80
Db	204	TTGGGAGAAGACATACATCTCCCAACTGAAGTTCATTTAAAAAACAGGTGGGATTTGTTCT	263
Qy	81	Pro a a a a g n leu n s arg g u asn le g n arg le val a g n u a leu Ser	100
Db	264	CCAGACGCTCAACAGCTTCACAGAGAAATATTCACGAATAGTACAAAGAGCTCTTCT	323
Qy	101	Ala Ser g y val Ser Pro Ser Asp leu Ser Ala l le a Thr Thr leu y Pro g y leu	120
Db	324	GCACATGGAGTCTCTCAAGTACCTCTCAGCAATTCACATACATTAACCAAGAGACTT	383
Qy	121	Ala leu Ser leu G y val g y leu Ser Phe Ser leu n leu val g y g n leu y s y s	140
Db	384	GCTTTAAAGCTCGGAGTGGCTTATCATTTAGCTTACAGCTGTGAGACAGCTTAATAAAG	443
Qy	141	Pro phe le Pro le n s e w et G u a n a e Ala leu Thr l le arg leu Thr Asn y s	160

Db		444	CCATTCAATCCCATTCATCATATGAGGCTCATGACTTACTATTAGGTGACCAATAA	503
OY		161	ValGluPheProPheLeuValLeuLeuIleSerGlyGlyHisCybLeuLeuAlaLeuVal	180
Db		504	GTAAAGATTTCCCTTTTATGTTCTTTTGATTTCTGAGGCTCACTGCTGTGGCATTTAGTT	563
OY		181	GlnGlyValIleSerAPheLeuLeuLeuGlyLyssSerLeuApriIleAlaProGlyAspMet	200
Db		564	CAAGGAGTTTCAGATTTTCTGCTTTTGAGAAAGTTTGGACATATGCACCAAGTACATG	623
OY		201	LeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGluCySerThrMetSer	220
Db		624	CTTGACAAAGTGGGCAAGAAAGCTTTCTTTAATTAACATCCAGAGTGTCCACCATGAGT	683
OY		221	GlyGlyLysValAlaIleGluHisLeuAlaLysGlnGlyAsnArgPheHisPheAspIleLys	240
Db		684	GGTGGGAAGCCATAGACATTTGGCCCAACAGAAATAGATTTTCATTTTGACATCTAA	743
OY		241	ProProLeuHisHISAlaLysAsnCybAspPheSerPheThrGlyLeuGlnHisValThr	260
Db		744	CTCCCTCTGCATCATGCTTAAATATGTGAATTTTCTTTTACTGCACTTCAACACGTTACT	803
OY		261	AspLysIleIleMetLysLysGlyLysGlyGluGlyIle-----	273
Db		804	GATAAAAATATATCAAAAAGGAAAAAGAGAGATATTTCTAATTAGTAAAGTTGA	863
OY		274	-----GluLys	275
Db		864	CAGATAAATATTCCTGATGATGCTTAAATAATAGCTGCTCATTTTTCGACGATAGAGAAG	923
OY		276	GlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetLysCys	295
Db		924	GGGCAAAATCCCTGCTTCAGACAGACACATGCTGCACAGTACAGACACACAAATGGCATGT	983
OY		296	HisLeuValLysArgThrHisArgAlaIleLeuPheCybLysGlnArgAspLeuLeuPro	315
Db		984	CATCTTGGAAGAAGACACATCGGCTATTCGTTTGTGTAAGACAGAGACTGTGTTACT	1044
OY		316	GlnAsnAsnAlaValLeuValAlaSerGlyValAlaIleSerAsnPheTrileArgArg	335
Db		1044	CAAAATATATGCAAGTACTGGTTCGACTCGGTGGTGTGCGCAATGACTTCTATATCGCAGAA	1103
OY		336	AlaLeuGluIleLeuThrAsnAlaThrGlnCybThrLeuLeuCybProProProArgLeu	355
Db		1104	GCTCTGGAATTTTAAACAAAGCAACACAGTCACTTGTGTGTCCTCTCCAGACTTA	1166
OY		356	CysThrAspAsnGlyIleMetIleAlaTPraGlnGlyIleGluArgLeuArgAlaGlyLeu	375
Db		1164	TGCATCTGAATATGCAATTATGATTTGCATGGAATGATTTGAAGAAGCTAGCTGTGGCTTG	1222
OY		376	GlyIleLeuHisAspIleGlyGlyIleArgTrgIuProLysCybAspProLeuGlyValAsp	395
Db		1224	GGCATTTTACATGACATAGAAAGCATTCGCTATGAAACCAAAATGTCTCTTGAAGTAAGC	1283
OY		396	IleSerLysGlyValGlyGluAlaSerIleLysValProGlnLeuLysMetGluIle	414
Db		1284	ATATCAAAAGAGTTGAGAAAGCTTCCATPAAAAGTACCAATTTAAAATGAGAGATA	1340

RESULT 3
 US-09-774-528-177
 ; Sequence 177, Application US/09774528
 ; Patent No. 6743619
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Goodrich, Kyle
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Ken, Feiyao
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Yang, Yonghong

```

APPLICANT: Xue, Aidong J.
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Wang, Dunrui
APPLICANT: Dymnac, Radoje T.
TITLE OF INVENTION: No. 674319e1 Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 802
CURRENT APPLICATION NUMBER: US/09/774,528
CURRENT FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 441
SOFTWARE: pc_fl_genes Version 2.0
SEQ ID NO 177
LENGTH: 1416
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (205)..(1305)
US-09-774-528-177

Alignment Scores:
Pred. NO.: 6.45e-213
Score: 1747.00
Percent Similarity: 100.00%
Best Local Similarity: 99.71%
Query Match: 82.21%
DB: 4
Length: 1416
Matches: 340
Conservative: 1
Mismatches: 0
Indels: 0
Gaps: 0
US-10-649-273-2 (1-414) x US-09-774-528-177 (1-1416)

```

OY	7	lyrthcglvcyllyleValaProproalaalainglnleuhi;argxlubasnl;eqlnarq	93
Db	280	AGAAcAGGTGGAGTTGTCTCTCCAGAGCTCAACAGCTTCAcAGAGAAATATTCACGA	333
OY	94	lleValcnglnuhalaleuSeralaSerGIyValSerProSeraspheuSerallleAla	113
Db	340	ATAGTACAAAGAGCTCTTCTCGCAATGAGTCTCTCAAGAGAGCTCTCGACCAATTGCA	399
OY	114	ThrThrlllelySerProGlyleuValaleuSerleuGIyValGIyleuSerPhaSerleuGln	133
Db	400	ACTACCAATAAAACAGAGACTGCTTTAAcCTGGAGTGGAGTGGCTTATCAATTAGCTTAAG	455
OY	134	leuValGIyGlnleuLylySerProPheileProleu;hi;emeGIuAlahi;salaleu	155
Db	460	CTGTAGAGAAcGTTAAAAAGcCATTCATTCcCATTCATCATATGAGGCTCATGCACTT	511
OY	154	Thrllleargleuthr;asnlyValGIupheProPheleuValleu;leu;leSerGIy	177
Db	520	ACTATTAGTTGAcCAATAAAGTAGAATTCTTTTATGTTCTTTGATTTCTGAGAGT	578
OY	174	hi;Scysleu;leuAlaleuValcnglyValSerasphe;leu;leu;leuGIylySerleu	193
Db	580	CACGTCTCTGTGGcATTAGTTCAAGAGATTTCAAGATTCTCTCTCTGGAAAGCTTTG	633
OY	194	Aspilleal;arProGIyaspmetleu;asp;lyValAla;arg;ArgleuSerleu;leu;hi;S	213
Db	640	GACATTAcACcAGGTGcCATGCTTGAcAAAGTGGCAAGAAAGCTTTCTTAATATAACAT	699
OY	214	ProGlyuSerSerThrm;eSerGIyGIylyValalleGlnhi;leuAlaleuGIcngly;asn	233
Db	700	CCAGAGTCTCCACATCATGAGTGTGGGAAGCATAGAAcATTGGCCAAACAGGAAT	755
OY	234	ArgPhehi;sphe;aplllelySerProPoleu;hi;hi;salaleu;asnCyaspPheSerPhe	255
Db	760	AGATTTCATTTCAGATCAAAcCTCCCTTCGATCATGCTAAAAAATGTGATATTTCTTTT	811
OY	254	ThrGIyleu;lnhi;val;Thr;asp;ly;si;leileme;ly;leuGIy;sgln;gln;glyle	277
Db	820	ACTGGACCTTCAACcCTTACTGATTAATAATATAAAAAAGCAAAAGAAAGAGAGTAT	877
OY	274	Gluylsgl;Gln;lleuSerSerAla;aspilleAla;leThrValcGlnhi;thrmet	293

Db 880 GAGAGGGGCAAACTCTGCTTCAAGCAGCAGCACTTGGCCACAGTACAGACACAAATG 939
Qy 294 AlaCysHisLeuValIysArgThrHisArgAlaIleLeuPheCysIysGlnArgAspLeu 313
Db 940 GCATGTCATCTTGTGAAAGAAACACATCGGGCTATCTGTTTTGTAGCAGAGACTTG 999
Qy 314 LeuProGlnAsnAsnAlaValIleuValAlaSerGlyGlyValAlaSerAsnPheTyrIle 333
Db 1000 TTAACCTCAAAATATATGAGTACTGTTGACATCTGGTGGTGTGCAAGTAACTTCAATATC 1059
Qy 334 ArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProPro 353
Db 1060 CGAGAGCTCTGGAATTTTAAACAAAGCAACAGACGACCTTGTGTGTCTCTCC 1119
Qy 354 ArgLeuCysThrAspAsnGlyIleMetIleAlaTrpAsnGlyIleGlnArgLeuArgAla 373
Db 1120 AGCATATGCACTGATTAATGGCATTATATGATGCAATGATGTAATGAAAGCTACGTCT 1179
Qy 374 GlyLeuGlyIleLeuHisAspIleGlyIleArgTyrGluProIysCysProLeuGly 393
Db 1180 GGGCTGGGCAATTTTACATGACATGAGAGGATCGGCTATGAAACCAAAATGCTCTTGA 1239
Qy 394 ValAspIleSerIysGlyValGlyValAspIleIysValProGlnLeuIysMetGlu 413
Db 1240 GTAGACATATCAAAAGAAAGTTGAGAGCTTCATTAAGTACCAATTAATAATGAG 1299
Qy 414 Ile 414
Db 1300 ATA 1302
RESULT 4
US-10-067-443-23
; Sequence 23, Application US/10067443
; Patent No. 6642041
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE HIGHLY EXPRESSED IN
; TITLE OF INVENTION: SPINAL CORD, MP-1
; FILE REFERENCE: D0073 NP
; CURRENT APPLICATION NUMBER: US/10/067,443
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 1526
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-067-443-23
Alignment Scores:
Pred. No.: 1,62e-166 Length: 1526
Score: 1385.00 Matches: 267
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.18% Indels: 0
DB: 4 Gaps: 0
US-10-649-273-2 (1-414) x US-10-067-443-23 (1-1526)
Qy 148 MecGlnAlaHisAlaLeuThrIleArgLeuThrAsnIysValGlnPheProPheLeuVal 167
Db 1 ATGAGAGGCTCATCACTTACTATATAGGTGACCAATTAAGTAGAATTTCTTTTACTT 60
Qy 168 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 187
Db 61 CTTTATATTTCTGAGAGTCACTGCTGTGGCATTAGTCAAGAGTTTCAGATTTTCTG 120
Qy 188 LeuLeuGlyIysSerLeuAspIleAlaProGlyAspMetLeuAspIysValAlaArgArg 207

Db 121 CTTTGGAAAGCTTTTGACATATGACACAGGTGACATGCTTGACAGGTGCAAGAGA 180
Qy 208 LeuSerLeuIleIysHisProGluCysSerThrMetSerGlyGlyLeuAlaIleGluHis 227
Db 181 CTTTCTTTATATTAACATCCAGAGTCTCCACCAAGAGTGGGGAAGCCATAGAACT 240
Qy 228 LeuAlaIysGlnIysAsnArgPheHisPheAspIleIysProProLeuHisHisAlaIys 247
Db 241 TTGGCAAAACAGAAATATGATTTTATTTGATCAATCAACCTCCCTGCATCATGCTAA 300
Qy 248 AsnCysAspPheSerPheThrGlyLeuGlnHisValIleThrAspIleIleIleLeuIys 267
Db 301 AATGTGATTTTCTTTTATCTGACCTTCAACAGTTACTGATTAATAATATGAAAAAG 360
Qy 268 GlyIysGlnGluGlyIleGluIysGlyGlnIleLeuSerSerAlaAlaAspIleAlaIa 287
Db 361 GAAAGAGAGAGATTTGAGAGGGGCAATCTGTCTTCAAGCAGACATGCTGCT 420
Qy 288 ThrValGlnHisThrMetAlaCysHisLeuValIysArgThrHisArgAlaIleLeuPhe 307
Db 421 ACAGTACAGCACACATGCGATGTCATCTTGGAAGAAAGACATCGGGCTATTTCTGTT 480
Qy 308 CysIysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyVal 327
Db 481 TGTAAAGAGAGAGCTGTGTTACTTCAAAATATATGAGTACGTGTGCATCTGTGTGTC 540
Qy 328 AlaSerAsnPheTyrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThr 347
Db 541 GCAAGTAACTTATATATCGGAGAGCTCTGGAATTTTAAACAAAGCAACACAGTGCAT 600
Qy 348 LeuLeuCysProProProArgLeuCysThrAspAsnGlyIleMetIleAlaTrpAsnGly 367
Db 601 TTGTTGTCTCTCTCCAGACTATGCACTGATTAATGATATATGATGCAATGAT 660
Qy 368 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGlyGlyIleArgTyrGlu 387
Db 661 ATGGAAGACATACGTGCTGGCTGGCAATTTTACATGACATGAGAGGCTACGTATGAA 720
Qy 388 ProIysCysProLeuGlyValAspIleSerIysGlyValGlyValAspIleIysVal 407
Db 721 CCAAAATGCTCTTGGAGTACATATCAAAAGAGTTGAGAGCTTCATTAAGTA 780
Qy 408 ProGlnLeuIysMetGluIle 414
Db 781 CCACATTAATAATGACATTA 801
RESULT 5
US-10-067-443-20
; Sequence 20, Application US/10067443
; Patent No. 6642041
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALOPROTEASE HIGHLY EXPRESSED IN
; TITLE OF INVENTION: SPINAL CORD, MP-1
; FILE REFERENCE: D0073 NP
; CURRENT APPLICATION NUMBER: US/10/067,443
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/266,518
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/282,814
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 14364
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-067-443-20
Alignment Scores:
Pred. No.: 2.19e-139 Length: 14364
Score: 1186.50 Matches: 313
Percent Similarity: 32.30% Conservative: 1

Best Local Similarity: 32.20% Mismatches: 2
Query Match: 55.84% Indels: 657
DB: 4 Gaps: 4
US-10-649-273-2 (1-414) x US-10-067-443-20 (1-14364)

QY 74 LysThrGlyGlyIleValProProAlaAlaGlnGlnLeuHISarGluAsnIleGlnArg 93
DB 10623 AGAAGCGTGGATTTGTTCTCCGACGCTCAACAGCTTCACAGAAAATATTCAGCA 10682

QY 94 IleValGlnGlnAlaLeuSerAlaSerGlyValSerProSerAspLeuSerAlaIleAla 113
DB 10683 ATAGTCAGAGAACTCTTCTGCGAGTGCCTCCAGTGCCTCCAGTGCCTCCAGCAATTCGA 10742

QY 114 ThrThrIleLysProGlyLeuAlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGln 133
DB 10743 ACTACCATTAACACGAGCTTCTGTTTACCTGGAGTGGGCTTATCATTTAGCTTACAG 10802

QY 134 LeuValGlyGlnLeuLysLysProPheIleProIleHISarGluAlaHISarAlaLeu 153
DB 10803 CTGGTAGACAGTTAAABAAAGCCATTCATCCCATCATATGAGAGGCTCATGCCTT 10862

QY 154 ThrIleArgLeuThrAsnLysValGluPheProPheLeuValLeuLeuIleSerGlyGly 173
DB 10863 ACTATTAGGTGGACCAATAAAGTAGAATTCCTTTTATGATTCCTTTGATTTCTGAGGT 10922

QY 174 HISarLeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeu 193
DB 10923 CACTGCTGTGGCATTAAGTTCAGAGAGTTCAGATTTTCGCTTCATGAGAAAGTCTTG 10982

QY 194 AspIleAlaProGlyAspMetLeuAspLysVal----- 204
DB 10983 GACATGACACGAGTGCATGCTTGACAGAGT-AAATTAAGATTAATTTCTCCATTCCTT 11041

QY 204 ----- 204
DB 11042 TTGTATGTTGCCATTTCACTAAGTAGCAATGATGCTACCAACATTCCTAA 11101

QY 204 ----- 204
DB 11102 TATTTCTGAATTTATCTTAGTAATACTGAAAAAATTCACATATGTCGAGAAAAATAGA 11161

QY 204 ----- 204
DB 11162 AAGAGTAGTACAAATTTATTTATTTAGCTTTCTTAATAAATGCTAAGAGTTTCT 11221

QY 204 ----- 204
DB 11222 ATCTGTACATTAAGGCTGAATACTGTCAGATACAGTATGATTTTGCCTAATAATGT 11281

QY 204 ----- 204
DB 11282 ATGTGAAGAAGCTTCTGTAACATACTGCAAAAAAGGTAAATTAAGAAATAT 11341

QY 204 ----- 204
DB 11342 ATATAGATTAAACATTAAGACATTAAGAATGCAAGATTAATCAACAATTAATCT 11401

QY 204 ----- 204
DB 11402 TACACCAACAGAGGTCCTCCCTCTTTGTTTGAATACTACAGAGGTAATCT 11461

QY 204 ----- 204
DB 11462 GCCATATATGAAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 11521

QY 204 ----- 204
DB 11522 TAGAGATATAGCAAGATCTTATTAATTAATGATTCATTAAGAGAGTATTTGCAACT 11581

QY 204 ----- 204
DB 11582 TCAAGCCCATTTTCCAACCAATAGAGCAAAACATAGACAGGGGCAAGTGAATGGCTCT 11641

QY 204 ----- 204
DB 11642 TATTTGCGGTCACTAATAGAAAGGGTGTCTGTTACTGTAATATCAGTAATAGTC 11701

QY 204 ----- 204
DB 11702 TATATTTGCCAAAGTATAGCATGTTTATTCATTCAGGGGTTTTTTGTTTGTAGTAAT 11761

QY 204 ----- 204
DB 11762 TTTCAATTTATTTCTTTCGATCTTTTCGTTTCACAGATTTAATTTTATGACTTAAA 11821

QY 205 ----- 205
DB 11822 AATATGTTTCTTGATAGGTGGCAAGAGACTTTCTTAATAAATCAATCCAGAGGCTCC 11881

QY 218 ThrMetSerGlyGlyLysAlaIleGlnHISarLeuAlaLysGlnGlyAsnArgPheHISarPhe 237
DB 11882 ACCATGAGTGTGGAAAGCCATAGAACATTTGGCCAAACAGGAATAATGATTTTCATTTT 11941

QY 238 AspIleLysProProLeuHISarAlaLysAsnCyAspPheSerPheThrGlyLeuGln 257
DB 11942 GACATCAAACTCCCTTGATCATAGCTTAATAATGTGATTTTCTTTACTGGACTTCA 12001

QY 258 HISarValThrAspLysIleIleMetLysGluLysGluGlu----- 271
DB 12002 CACGTTACTGATAAATAATATAGAAAAAGGAAGGATATTTCTAATTAAGT 12061

QY 272 ----- 272
DB 12062 AAGTTGAACAGATAAATATCTCGATGTCCTAAATAATAGCTCATATTTCTGCAG 12121

QY 272 YIleGlnLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaThrValGlnHISar 292
DB 12122 TATTAAGAAAGGGCAAACTCTGCTTCAGACAGACATTCCTGCACAGTACAGACAC 12181

QY 292 IMetAlaCyHISarLeuValLysArgThrHISarAlaIleLeuPheCysLysGlnArgAs 312
DB 12182 AATGCAATGATCTTTGTGAAAGAACACATCGGCTATTCTGTTTGTAAAGAGAGCA 12241

QY 312 PLeuLeuProGlnAsnAsnAlaValLeu----- 321
DB 12242 CTTGTTACTCAAAATATATGAGTACTGTAAAGTTTATCTCATTTATAGTAATAGTTA 12301

QY 321 ----- 321
DB 12302 CACTTGCAATATGTACTTTTTCCTCAAGACTTGACTGTGTGTTAGATGAACAGAT 12361

QY 321 ----- 321
DB 12362 CTTTATGCTTAATGCTACCCCTGACAGATTAATAATTAAGAGATAGAAAGCTAACAG 12421

QY 321 ----- 321
DB 12422 CCATTTCTTGACTAGTTTGTAGCTTTATAGGACAGCTGTATAGCTTCTATAGCACATA 12481

QY 321 ----- 321
DB 12482 AGCTAATTTTGATCTTCTGTGTGATTTAAAGAGGCTTACAATTAAGAAAGTAAT 12541

QY 321 ----- 321
DB 12542 GCAGTAAGCTGTATCACTAATTTTGAAGAAAAATAGGTGATTTCTTCATCTTTGATGA 12601

QY 321 ----- 321
DB 12602 ATCCCTTGTGTTGTTGTTTATTAATAGCAAGTCAATTTAGCAGTGGAGGTGAT 12661

QY 321 ----- 321
DB 12662 TCCAACTTGCGTGAACATAATGTTGATAAAGTTCTGATTAATCACTAATATTTGACAGCC 12721

```
QY 321 ----- 321
Db 12722 AAAATCCCTTAATGCTTAATAAGCCTTGACAAACATCCCTTTAACTGATCTTAAA 12781
QY 321 ----- 321
Db 12782 CTTTATTCATTTAAAAATTATATACTAAGTGGAAAAATTTAATGTAGTAATTCAT 12841
QY 321 ----- 321
Db 12842 AGATGGAATTTTACATGATATCAAGAAATATTTTTCAGAGTATGTAGTAAATGCA 12901
QY 321 ----- 321
Db 12902 CAAAATAATAAAATTTTACGGGTCTAAATAGTGTACTATGATGAATTAATTAATA 12961
QY 321 ----- 321
Db 12962 AATATTAGATGAAGGTGGAGAAAAATATACAAAAATGCTAGTAATGTTGTATGCTA 13021
QY 321 ----- 321
Db 13022 TTAGATTAATTAATTTTCTTCCAAATTTTATTAACATAGATATGTCATCTGCC 13081
QY 321 ----- 321
Db 13082 CATTAACCATCTCAAAATGGAGTAGTTATTAATGTTAATGCTGATATTTTCTCAGG 13141
QY 321 ----- 321
Db 13142 TTTAATTAGCAGCTGGTTCATATCCATATATGATAGTATTTGTTTCTCAATTCCT 13201
QY 322 -----ValAlaSerGlyValAlaSerAspPheTyrIleArgArgAlaLeuGluIle 340
Db 13202 TCAGGTGCACTTGCGGTGTCGACAGTAACTTCTATATCCGACAGCTCTGGAAATTT 13261
QY 340 uThrAsnaIaIhGlnCysThrLeuLeuCysProProArgLeuCysThrAspAsnG 360
Db 13262 AACAAACGACACAGTGCATCTTGTGTCTCCCTCCAGACTAGTACCTGATATGG 13321
QY 360 yIleWcTlleAlaTTP----- 365
Db 13322 CATTAATGATTCAGTGAAGCCACAGAGTATACGTGCTCATCATTAATGTAATAT 13381
QY 365 ----- 365
Db 13382 TAATTGCCATTTTATCATTAAGCCTTCTTCTTCAGATCTTGAGCTATGATTTTAT 13441
QY 366 -----AsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuH 379
Db 13442 TTTAATGCTTCTTATTTAGAAATGCTATTAAGAACTAAGCTGCTGGCATTTTAC 13501
QY 379 IAspIleGluGlyIleArgTyrGluProLys 389
Db 13502 ATGACATAGAAAGCATCCGCTATGAACCAAG 13553

RESULT 6
US-09-596-002-38
; Sequence 38, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Paterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 38
```

```
; LENGTH: 94750
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURES:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyre template ID No. 6632636 38
; PUBLICATION INFORMATION:
US-09-596-002-38

Alignment Scores:
Pred. No.: 2,66-50 Length: 94750
Score: 502.00 Matches: 129
Percent Similarity: 50.81% Conservative: 59
Best Local Similarity: 34.86% Mismatches: 149
Query Match: 23.62% Indels: 34
DB: 4 Gaps: 9

US-10-649-273-2 (1-414) x US-09-596-002-38 (1-94750)
QY 39 ValLeuGlyIleGluThrSerCysAspAspThrAlaAlaValAlaAspGluThr--- 57
Db 23873 GTATTGGATTGGAGACATCTTGATGAAACAGGGCTGCCATCTATGATAGTACAGTG 23932
QY 58 -----GlyAsnValLeuGlyGluAlaIleHisSerGlnThrGluValHisLeu 73
Db 23933 AATGGCGGTGCGCGCGCGCTGCTCAGGTTTGTATACCAAAATTAATCATATGCC 23992
QY 74 LysThrGlyIleValProProAlaGlnGlnLeuHisArgGluAsnIleGlnArg 93
Db 23993 ACTATAGCGGTGCGCGCGCTGACCTTGCCAGTGGAGACCAATTCGTAACCTGTGCGG 24052
QY 94 IleValGlnGluAlaLeuSerAlaSerGlyValSerProSerAspLeuSerAlaIleAla 113
Db 24053 TTATTTAATAGCTGTTGATGATCAGGCAAAATATCAACCAATCCGATGATGCGGTGCG 24112
QY 114 ThrThrIleLysProGlyLeuAlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGln 133
Db 24113 TATACCAAGAGGCGCAGGCGATTTGGGGCGTGTATGACAGGGCGATTTTGGGCGAG 24172
QY 134 LeuValGlyLeuLysLysProPheIleProIleHisIleMetGluAlaHisAlaLeu 153
Db 24173 CTGGCGTATGGCGGTGCGCGTGCAGCGTGGGGTGTATGAGGAGCATCTGTTA 24232
QY 154 ThrIleArgLeuThr-----AsnLysValGluPheProPheLeuValLeuLeuIleSer 171
Db 24233 GCACCGCTATTTGGCCAGTATGACCCCAAGCTTTCGTTGTGTGCTCTGCTGCTGCTG 24292
QY 172 GlyIleHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLys 191
Db 24293 GCGGTATACCAACGCTGTCGTCGCGATGCTGTGGCGCGTATCAGATTTGGGCGAG 24352
QY 192 SerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIle 211
Db 24353 TCTATCGATGATGCGGTGCGTGAATGCTTGATTAACCGCAAAATGCTCAAACTG--- 24409
QY 212 LysHisProGluCysSerThrMetSerGlyGlyLysAlaIleGlnHisLeuAlaLysGln 231
Db 24410 CCTATCT-----GTTGGCCCAATATGCAAAATTAATGCAAAAC 24451
QY 232 GlyAsnArgPheHisPheAspIleLysProProLeuHisIleAlaLysAspCysAspPhe 251
Db 24452 GGCACCCACACCGCTATGAGCTGCCAAGACCAATGACGAT---AAAGGCTGATTTT 24508
QY 252 SerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLysGluLysGlu 271
Db 24509 TCGTTTCATGCGATGAAGAACCGCATTCATATCTCATCAAGACACACCAAGGCCCAA 24568
QY 272 GlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaIleThrValGlnHis 291
Db 24569 AGCGACCC-----GCCACACGACAGACATCGCGCAAGCTTTGAGAT 24613
QY 292 ThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPheCysLeuGlnArg 311
DB: 4 Gaps: 9
```

```

Db      24614 CCGGTGGTGGATACCTTTGGTCAAAAATATGACCAAGACACTACAGATGACAGGACTTCGC 24673
Qy      312 AspleuleuProGlnAsnAsnAlaValIleuValAlaSerGlyValAlaSerAsnPro 331
Db      24674 CAG-----CTGTGTGTGGCAGAGGGGGCGTCTCTCCACTACAG 24709
Qy      332 TyrIleArgArgAlaIleuGluIleLeuThrAsnAlaThrGlnCysThrIleuLeuCysPro 351
Db      24710 ATGCTACGCCGACCCCTGACCGAGACGCTCCGCCAATATGATGCGTGGGTGACTATGCC 24765
Qy      352 ProProArgLeuCysThrAspAsnGlyIleMetIleAlaIleArgLeu 371
Db      24770 CCGACCGACACTATGACGAGATATGTGTGCAATGCTATGCTGGCTTTGTGGGCTC 24829
Qy      372 ArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGluProLysCysPro 391
Db      24830 AGCTGTGGACAGTCG-----GATGACTTGTGGC-GGTTGCTGTATTC-----CCG 24873
Qy      392 LeuGlyValAspIleSerLysGluValGly 401
Db      24874 ATGGGATATGACGACGCTTGCGGTATCGGC 24903

RESULT 7
US-09-540-236-806
; Sequence 806, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540.236
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 806
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: M. catarrhalis
US-09-540-236-806

```

Alignment Scores:	1.23e-53	Length:	1053
Pred. No.:	501.00	Matches:	126
Score:	49.86%	Conservative:	56
Percent Similarity:	34.52%	Mismatch:	139
Best Local Similarity:	23.58%	Indels:	44
Query Match:	4	Gaps:	8

US-10-649-273-2 (1-414) x US-09-540-236-806 (1-1053)

```

QY      39  VALLLEUGLYILEGIUTHRSERYASPAAPHTRALAAALVALVALSPGIUTHR--- 57
          |||:::|||||:::|||||:::|||||:::|||||
Db      16  GATATGGGATGTGAGACATCTTGATGAAACAGGCGCTGCATCATATATATAGCATG 75
QY      58  -----GLYASNAVLEUGLYGLUALALHHSISERGLTHCLUVALHISLEUYLS 74
          |||:::|||||:::|||||:::|||||:::|||||
Db      76  AATGGCAGGGGGCGCGCTGCTCTCAAGTTTATATAGCCAAATPATATCATACGCCACC 135
QY      75  THNGLYGLYILEVALPROFOLALAGLNGLNLEWHISARGLUANILLEGIMARGILE 94
          |||:::|||||:::|||||:::|||||:::|||||
Db      136  TATGGCGGATCGTCCGCTGAGCTTGCCCATGACAGACCACATCGTAACTGTGCGCGTGA 195
QY      95  VALGNGLUALALEUSERRALASERCLYVALSERPROSERASPUSERSALALALATHR 114
          |||:::|||||:::|||||:::|||||:::|||||
Db      196  TTTAATGAGATGTTGATCAGGCAAAATATCACAAATCCGACATTTAGTGGTGCGGTAT 255
QY      115  THNILEYSPROGLYLEUALALEUSERLEUGLYVALGILEUSERPHESERTLEUGINLEU 134
          |||:::|||||:::|||||:::|||||:::|||||
Db      256  ACCAAMAGGCCCGCGCGCATTTGGGGCAATTGATGACAGGGGCAATTATTTGGCGGACGCTG 315
QY      135  VALGILEULEUYLSYSPROPHETLEPROLEHSHIMETGLUALHSHALALAUYTHR 154
          |||:::|||||:::|||||:::|||||:::|||||
Db      316  GCGTATAGCGTGGCGCGTGGCCAGCGGTGGGGCATCATATATGAGGACATCTGTTAGCA 375

```

[illegible]

```

RESULT 8
US-09-252-991A-801
; Sequence 801, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS.
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 801
; LENGTH: 1206
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa

```

US-09-252-991A-801

Alignment Scores:

Pred. No.: 2,43e-53 Length: 1206
 Score: 499.50 Matches: 133
 Percent Similarity: 51.68% Conservative: 67
 Best Local Similarity: 34.37% Mismatches: 158
 Query Match: 23.51% Indels: 29
 Gaps: 10

US-10-649-273-2 (1-414) x US-09-252-991A-801 (1-1206)

```

OY 13 LysProSerLeuValGlyValThrLeuValSerPheAsnProGlyThr 32
DB 91 AAAGCCTCCGATCGAAGAACTGCTGCGCGAAGCGCGCTTATATACGCGGCTT 150
OY 33 LeuPheLeuHisLysLe-----ValLeuGlyLeuGlyThrSerCys 46
DB 151 CGTATTGCGAAGCTTAAAGCTCAAGCCCAATCGCGCTGCGAAGCGCTTGC 210
OY 47 AsnAspThrAlaAlaAlaValAlaValAspGlyThrGlyAsnValLeuGlyAlaLeuHis 66
DB 211 GACGAAACCGCGGCTGCTTATACGACGAAAGCGGCTGCTGCGCGAAGCGCTTTC 270
OY 67 SerGlnThrGlnValHisLeuLysThrGlyGlyLeuValProProAlaAlaGlnGlnLeu 86
DB 271 AGTCAGATGACCTCATCGCTGACGCGGCGGCTGCTGCGCGAAGCTTGC 330
OY 87 HisArgGlnAsnHisLeuGlnArgLeuValGlnGlnAlaLeuSerLysAspGlyValSerPro 106
DB 331 CACGTCAAGCGATGCTGCGCGCTGATCCGCGAGCTGCTCAACGATCCGCTGACGCGC 390
OY 107 SerAspLeuSerAlaLeuAlaThrThrThrLeuLysProGlyLeuAlaLeuSerLeuGlyVal 126
DB 391 GCGGATATGACGCGATCCGCTATACCGCGCGCTGCTGCGCGCTGCTGCTGCTGCTG 450
OY 127 GlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuLysLysProPheLeuProLeuHis 146
DB 451 GGGGCTTCTGTGCGCGAGCGATGCTTGCCTGCGCGCTGCGCGCGCTGCGCGCTG 510
OY 147 HisMetGlnAlaHisAlaLeuThrThrLeuGlyLeuThrAsnLys---ValGlnPheProPhe 165
DB 511 CACATGGAAGGCGACCTGCGCGCGATGCTGGAAGAGACACCGCGCTTCCCGCTTC 570
OY 166 LeuValLeuLeuLeuSerGlyGlnHisCysLeuLeuAlaLeuValGlnGlyValSerAsp 185
DB 571 GTCGCTTCTGCTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGC 630
OY 186 PheLeuLeuLeuGlyLysSerLeuAspLeuAlaProGlyAspMetLeuAspLysValAla 205
DB 631 TACCAAGTTGCTGCGCGAATCGGTGAGCGATGCGCGCGCGAAGCGCTTCCGACAGACCGCC 690
OY 206 ArgAlaGlyLeuSerLeuLeuHisHisProGlyCysSerThrMetSerGlyGlyValAlaLe 225
DB 691 AAGCTGATCGCGCTG---GGCTATCC-----GGTGTCCGGAATC 729
OY 226 GlnHisLeuAlaLysGlnGlyAsnArgPheHisPheAspLeuLysProProLeuHisHis 245
DB 730 GCCCGCTGCGAGCGCGCACTCTGCGCGCTGCTGCTGCTGCTGCGCGCGAATCGCGAT 789
OY 246 AlaLysAsnCysAspPheSerPheThrGlyLeuGlnHisValAlaThrAspLysIleIleMet 265
DB 790 CGCCCGCGCTGACCTTCACTTCAAGCGGCTCAAGACCTTACCTTCAAGAC---ACCTGG 846
OY 266 LysLysGlnLysGlnGlnGlyIleGlyLysGlyGlnIleLeuSerSerAlaAlaAspIle 285
DB 847 CAGCGTGTGCTGAGCGCGCGCGACGACGAGCAG-----ACCGCTCGACATC 897
OY 286 AlaAlaThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisAlaAlaIle 305
DB 898 GCGCTGCGCTTCCAGACCGCGGTGTCGAGACCTCTGATCAAGCGCGCTGCGCGCTG 957
OY 306 LeuPheCysLysGlnArgAspLeuLeuProGlnAsnAlaValAlaLeuValAlaSerGly 325
  
```

```

DB 958 -----AAGCAGCGGCTG-----AAGAAC-----CTGTATGCGCGC 993
OY 326 GlyValAlaSerAsnPheThrLeuArgAlaLeuGlnIleLeuThrAsnAlaThrGln 345
DB 994 GGTGTACGCCCAACGAGCGGCTGCGAGGCGCTGGAAGAAAGTCGCGGAATAGAG 1053
OY 346 CysThrLeuLeuCysProProProAlaGlyCysThrAspAsnGlyIleMetIleAlaTrp 365
DB 1054 GGGCAGGTGTCTTACGCGCGCGCTGCTGACCGACATGCGCGCATGATGCGCTTAC 1113
OY 366 AsnGlyIleGlnArgLeuAlaGlyLeuGlnIleLeuHisAspIleGlnGlyLeuArg 385
DB 1114 GCGCGCTGCGCGCGCTGCTGCGCGC-----CAGCATGACGCGCGCGATTCAGC 1164
OY 386 TyrGlnProLysCysProLeu 392
DB 1165 GTCCAGCGCGCTGCGCGATG 1185
  
```

RESULT 9

```

US-09-252-991A-884/c
; Sequence 884, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 884
; LENGTH: 1059
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-884
  
```

Alignment Scores:

Pred. No.: 1.52e-52 Length: 1059
 Score: 492.50 Matches: 128
 Percent Similarity: 53.24% Conservative: 61
 Best Local Similarity: 36.06% Mismatches: 143
 Query Match: 23.18% Indels: 23
 Gaps: 9

US-10-649-273-2 (1-414) x US-09-252-991A-884 (1-1059)

```

OY 39 ValLeuGlyIleGlnThrSerCysAspAspThrAlaAlaValAlaAspGlyThrGly 58
DB 1038 GTGTGTGAGCTGGAAGCTCTTGAGCAAGAAACCGCGCTGCTTACGACGGAACGC 979
OY 59 AsnValLeuGlyGlnAlaLeuHisSerGlnThrGlnValHisLeuLysThrGlyGlyIle 78
DB 978 GGCCTGCTGCGCGCGCGCTTCAATGATGACATGCACTTCAATCGCGCTTACGCGGCTC 919
OY 79 ValProProAlaAlaGlnGlnLeuHisArgGlnAsnIleGlnArgIleValAlaGlnAla 98
DB 918 GTGCGGAGCGGCTGCGCGGACGACGTCAGAGCGATGCTGCGCGATGCGCGAGGTG 859
OY 99 LeuSerAlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleLysPro 118
DB 858 CTCGACGAGTCCGCTGCTGCGCGCGCGATATGACGCGATGCGCTTACCGCGCTTCC 799
OY 119 GlyLeuAlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeu 138
DB 798 GGCCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 739
OY 139 LysLysProPheLeuProLeuHisHisMetGlnAlaHisAlaLeuThrThrIleArgLeuThr 158
  
```


Db 946 ACCGATATATGTCCTTGGCGGAGTATCCGTTTAAAGTGTACCGAG 1005
Qy 376 GYILEUHIHISAPILLEGUJLYIARGYRGLUPROBLYSCYSPROLEU 392
Db 1006 GGGCCCTTTA-----GGGGTGAAGTGAAGACAGTTGGCCCTTTA 1044
RESULT 11
US-09-902-540-6612
; Sequence 6612, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PENDING FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 6612
; LENGTH: 996
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-6612
Alignment Scores:
Pred. No.: 3,01e-51 Length: 996
Score: 482.00 Matches: 128
Percent Similarity: 52.48% Conservative: 52
Best Local Similarity: 37.32% Mismatches: 131
Query Match: 22.68% Indels: 32
DB: 4 Gaps: 9
US-10-649-273-2 (1-414) x US-09-902-540-6612 (1-996)
Qy 38 ILEVALLEUGJLYILEGUTHSERCYSPASPPTHRAIAIAIAVAIAASPGLUTHR 57
Db 4 CTCGCTCTAGAGATGGAACCTGTGTATGAGACTCCGCGCGCTGTGAGAGAGAGCC 63
Qy 58 GIYASNAVLEUGJLYIAIAIAIEHISERGINTHRIGLYIAHISLEUYSRTHGLY 77
Db 64 CCGCGCGCTGTGCGATGTCGTCCAGCAGGTGACATCCACCGCGGTGGCGGG 123
Qy 78 ILEVALPROBIAIAIAGINTLEUHIHISARGIUNSLIEGINARGLILEVALGINLU 97
Db 124 GTGGTCCGAGCTGCGCAGCCCAACACATGCTCCAGTGTGCGCGTCCAGAG 183
Qy 98 AIALEUSERIASERGIYALSERPROSERASPLEUSERIAIAIAIAIAIAIAIA 117
Db 184 GGGCTACCGCGGCGCAACAGAGCTTCGACAGCTGACCTGACCTGACCTCGGC 243
Qy 118 PROGLYLEUAIALAEUSERLEUGJLYIAGLYLEUSERPHEUSERLEUGJLYIAG 137
Db 244 CCGGACTATCGCGCGCGCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 303
Qy 138 LEULYELSPROBIEHIEPROIEHISIMETGLUAIHISALAEUTHRIIEARGLEU 157
Db 304 ACGGCGAAGCCCTTCTGGGCGGCAACACCTGAGGCGACCTGCGCATCCGCTG 363
Qy 158 THRASLUYVAL---GLUPHEPROBIEUVALLEULEULESERGIYGLYHISYLEU 176
Db 364 TTGAGAGTGGCGCGGAGCGCGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 423
Qy 177 LEUAIALAEUVALGINGLYVALSERASPPHELEULEULESERGIYSSERLEUAI 196
Db 424 CTCCTACGAGGTGAGCGCTTACGCGGCACTGCTGTGTGAGCAACCGCGAGAGCC 483
Qy 197 PROGLYSPMETLEUASPLYSVALAIARGARGLEUSERLEULEYSHISPROGLUCYS 216
Db 484 GCGGCGAGAGCAATATGACAAAGACCGCTGATCTCGGCGCTG---CCGTATCCG----- 534

Qy 217 SERTHMETSERGIYLYSALAIIEGLUHIHISLEUALALYSGLINGLYASN----- 233
Db 535 -----GGTGGCGAGCCCATGACAGCTGTGGCGAGCGAGGAAACCGGAGGCC 582
Qy 234 ---ARGPHEHISPREHAPILIELYSPROBIEUHIHISIALYASPCYASPPHESE 252
Db 583 ATCCGCTTC-----CCGCGCGCGCTGCGCGGCGCAACACTTGCAGCTGTCC 627
Qy 253 PHEATHRIGLEUGJINHISVALTHRASPLYSILEIMETLYSLYGLUYGLUGJLY 272
Db 628 TTCTCCGGGTTGAAG-----ACGCGGTGCTGCACCACTGACAGACAGCGC 675
Qy 273 ILEGUJLYSGJLILIEUSERSERALAIASPLIEALAIATHRVAIGLNIHISTR 292
Db 676 GTGCGGAGGCGGAGGCGCTG-----GCGGATTTGTCGGGTCTTCAGAGAGGCC 726
Qy 293 METIACYSHISLEUVALYARGTHRIARGLAIALEULEUPHECYSLYGLINARGSP 312
Db 727 GTGCGGACGTCGTGTCAAGAGAG-----CTGTGGCCCGCGCGCGCGG 771
Qy 313 LEULEUPROGINASNAIAIAIAIAIAIAIAIAIAIAIAIAIAIAIAIAIAIAIA 332
Db 772 TTG-----GCCCAAGACAGTGTGTCTGTGCGCGCGCTGCGCGGACTCGCGG 822
Qy 333 ILEARGAIALEUGJLILIEUTHRASNAIAIAIAIAIAIAIAIAIAIAIAIAIA 352
Db 823 CTGCGGACAGTGTGTCAAGCGCGAGCGAGGCGGCGGTGAACATGTTCTCGCCCG 882
Qy 353 PROARGLEUCYTHRASPNGLYIIMETILEALATRPASNGJLYILEGUTHLEUARG 372
Db 883 GTGCGGCTGTGACGAGCAATATGCGCATATTCGCTGTGCGGCTATAGAGCGTACCGC 942
Qy 373 AIAGLYLEU 375
Db 943 CCGCGCGCTG 951
RESULT 12
US-09-902-540-503/C
; Sequence 503, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PENDING FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 503
; LENGTH: 2582
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-503
Alignment Scores:
Pred. No.: 1.62e-50 Length: 2582
Score: 482.00 Matches: 128
Percent Similarity: 52.48% Conservative: 52
Best Local Similarity: 37.32% Mismatches: 131
Query Match: 22.68% Indels: 32
DB: 4 Gaps: 9
US-10-649-273-2 (1-414) x US-09-902-540-503 (1-2582)
Qy 38 ILEVALLEUGJLYILEGUTHSERCYSPASPPTHRAIAIAIAIAIAIAIAIA 57
Db 1525 CTCGCTCTAGAGATGGAACCTGTGTATGAGACTCCGCGCGCTGTGAGAGAGAGCC 1466

```

Oy 58 GlyAsnValLeuGlyAlaIleHisSerGlnThrGluValHisLeuLysThrGlyGly 77
Db 1465 CCCCCGCGTTGGATGTCGTCACGACGATGACACGCGCGGTGGGG 1406
Oy 78 IleValProProAlaIleGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGlu 97
Db 1405 GTGGTCCCGAGCTGCGACGCGCACACATGTCGACGTCGCCCGCTCCACGAG 1346
Oy 98 AlaLeuSerAlaSerGlyValSerProSerAlaIleAlaThrThrIleLys 117
Db 1345 GCGCTACCGCGCGGACACAGACGCTTCGACGACGATGACCTCATCGCTACGCTCGGC 1286
Oy 118 ProGlyLeuAlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnValGlyGln 137
Db 1285 CCGGACTCATGCGCGCGCTGCGTGGAGTGCAGAGTGCACAGGCTTGAGCTGGCG 1226
Oy 138 LeuLysLysProPheIleProIleHisMetGluAlaHisAlaLeuThrIleArgLeu 157
Db 1225 ACCGACAGCCCTTCGTCGCGCGCACACCTGAGGCGCACCTGCGCATCCGCGCTG 1166
Oy 158 ThrAsnLysVal---GluPheProPheLeuValLeuLeuLysSerGlyHisCysLeu 176
Db 1165 TTGGAGCTGCGCGCGCGCGCTTCCTGCGCTGCTTCGCGCGCGGACACAGC 1106
Oy 177 LeuAlaLeuValGlnGlyValSerPheLeuLeuGlyLysSerLeuAspIleAla 196
Db 1105 CTCTACGAGTGCAGCGCTACCGGACGCTGCTGGGCGACGCGCGACGAGCGG 1046
Oy 197 ProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGluCys 216
Db 1045 GCGCGGAGCATATGACAGACCGCTCGCATCTCGCGCTG---CCGTATCCG----- 995
Oy 217 SerThrMetSerGlyGlyLysAlaIleGlnHisLeuAlaLysGlnGlyAsn----- 233
Db 994 -----GTTGGCAGCCCGCATGACAGCTTGGCGCAGCAGCGGGAACCCGAGAGCC 947
Oy 234 ---ArgPheHisPheAspIleLysProProLeuHisHisAlaLysAsnCysAspPheSer 252
Db 946 ATCCGCTTC-----CCGCGCGCTGCGCGGCGCACACTTCGACGTGCTC 902
Oy 253 PheThrGlyLeuGlnHisValThrAspLysIleIleMetLysGlyLysGlnGluGly 272
Db 901 TTCTCCGGGTGAG-----ACGCGGTGCTGCACACCGTCGACAGACGCGC 854
Oy 273 IleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaIleThrValGlnHisThr 292
Db 853 GTCCGCGAGGCGAGCGCGCTG-----GCGGATTTGTCGCGCTCTTCACGAGGCC 803
Oy 293 MetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAsp 312
Db 802 GTGCGGACGCTGCTGCGAAGAG-----CTGTCGCGCGCGCGCGCGCG 758
Oy 313 LeuLeuProGlnAsnAlaValAlaSerGlyValAlaSerGlyValAlaSerPheThr 332
Db 757 TTT-----GCCACAGACGTTGCTGTCGCGCGCGCTGCGCGGACATCCGCG 707
Oy 333 IleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLysProPro 352
Db 706 CTCGCGGACATGTCACGCGCGACGCGAGGAGCGGGGTGAACATGTTCTGCCCCG 647
Oy 353 ProArgLeuCysThrAspAsnGlyIleMetIleAlaThrAsnGlyIleGluArgLeuArg 372
Db 646 GTCCGCTGTGCACGACATGCGCGCATGTCGCGTGGCGGGGTATGAGCGCTACCGC 587
Oy 373 AlaGlyLeu 375
Db 586 CCGCGCGCTG 578

```

RESULT 13
 US-08-087-797-1
 ; Sequence 1, Application US/08087797
 ; Patent No. 5543312
 ; GENERAL INFORMATION:

```

APPLICANT: Mellors, Alan
APPLICANT: Lo, Reggie Y.C.
APPLICANT: Abdullah, Khalid M.
TITLE OF INVENTION: Pasteurella Haemolytica
TITLE OF INVENTION: Glycoprotease
TITLE OF INVENTION: Gene and the Purified Enzyme
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSER: Bell, Seltzer, Park & Gibson, P.A.
STREET: 1211 East Morehead Street,
CITY: Charlotte
STATE: No.5543312th Carolina
COUNTRY: United States
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,797
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Layton, Jr., Samuel G.
REGISTRATION NUMBER: 22807
REFERENCE/DOCKET NUMBER: 3374-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 704 377 1561
TELEFAX: 704 334 2014
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1315 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 141..1115
US-08-087-797-1
Alignment Scores:
Pred. No.: 1,19e-50 Length: 1315
Score: 479.00 Matches: 128
Percent Similarity: 48.63% Conservative: 50
Best Local Similarity: 34.97% Mismatches: 157
Query Match: 22.54% Indels: 31
Gaps: 8
US-10-649-273-2 (1-414) x US-08-087-797-1 (1-1315)
Oy 9 GlyValPhePheLysProSerLysArgLysValTyrGluPheLeuArgSerPheAsp 28
Db 89 GGGGCGTCTTCCGCCCT-----TTGGTTTTCACCTATTATTTGACTTC 133
Oy 29 HisProGlyThrLeuPheLeuHisLysIle-----ValLeuGlyIleGlnThrSerCysAspAs 48
Db 134 TCC-----AACATGCGAATTATAGTATGAAAGGCTTATGAAAGGCTTATGAGCA 175
Oy 48 PThrAlaAlaValAlaValAspGluThrGlyAsnValLeuGlyGluAlaIleHisSerG 68
Db 176 AACCGTGTGCAATTATGATGAGACAAAGGCTTAGCGGCAACAGCTTATAGCCA 235
Oy 68 PThrGluValHisLeuLysThrGlyGlyIleValProProAlaIleGlnGlnHisArg 88
Db 236 AATTGATATGACCGCGCATTAAGGTATGAGTATGCTGCTCACTGCTTCGACACCAT 295
Oy 88 GlnAsnIleGlnArgIleValGlnGluAlaLeuSerAlaSerGlyValSerProSerAs 108
Db 296 CCGTAAGAGCTTGCACATTAATTCAGAGGCTTAAGAGGCGCAATTCGACACCTGCGA 355

```

```
QY 108 pLeuSerAlaIleAlaThrThrIleIleYsProGlyLeuAlaLeuSerLeuGlyValGlyLe 128
Db 356 TATTAAGCGCATTTGCTTAACTATCCGCGCCAGCTTGTGCGGCTTATATGTCGCGCTC 415
QY 128 uSerPheSerLeuGlnLeuValGlyGlnLeuYsIleProPheIleProIleIleShiSme 148
Db 416 AACCAATGCGCCCTTCCCTGCTTATGCTTGAATGTTGCGGCTTGTGCGCTTATATG 475
QY 148 tGluAlaHisAlaLeuThrIleArgLeu---ThrAsnLysValGlyPheProPheLeuVal 167
Db 476 GGAAGGCGATTACTTCTGCGCAATGTTGGAAGAAATGCGCTGAAATTTCTGTTTGCGC 535
QY 167 lIleuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerPheLe 187
Db 536 ATTATTGATTTCAGGTCGACACACCAACTGTTAAAGTTGACGGGCTTGCGCAATACGA 595
QY 187 uLeuLeuGlyIleSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 207
Db 596 ACTACTCGGGGAAATCAATTGATGATGCTGCGGTGAAGCCTTGTGACAAACAGGCAAACT 655
QY 207 gLeuSerLeuIleLysHisProGlyCysSerThrMetSerGlyIleLysAlaIleGlu 227
Db 656 ACTCGGTTTG---GATTACCTT-----GCCGGTGTACGCAATGTCAAA 694
QY 227 sLeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisShiSala 247
Db 695 ATTACCGCAATCCGACCGCAAAATGCTTTAAATTCCTCCGTCATGACGACAGACACC 754
QY 247 sAsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLys 267
Db 755 GGGACTGGATTTTCAGTTTCTCCGCTTAAACCTTGTGCGCAATACATTAAGCCAA 814
QY 267 sGluLysGlnGlyGlyIleGlyLysGlnIleLeuSerSerAlaAlaAspIleAlaI 287
Db 815 TCTTAATGAATAATGTGAATCTCGATGACGAA-----ACCATAATGCCATATTGCGCA 865
QY 287 aThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeu 307
Db 866 CGCATTCACAAACAGCCGTC-----GTTGATACATATTTAATTA 904
QY 307 eCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyIle 327
Db 905 ATGCAAG---CGAGCGCTTAGAGCAAAACCGCTATTAACGCTTAGTATGACAGCGCGCT 961
QY 327 lAlaSerAsnPheThrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCys 347
Db 962 AAGTCCCAATTAACAATTACAGACAGACCTTGGCGAAATGATGAAAAAATTAAAGCGCA 1021
QY 347 rLeuLeuCysProProArgLeuCysThrAspAsnGlyIleMetIleAlaThrAsnG 367
Db 1022 AGTATTCTTACCTTCGCCCAATTTTGCATGACACGCGCAATGTTGCTTACACTGG 1081
QY 367 yIleGluArgLeuArg 372
Db 1082 CTTTCTGCTTAAAA 1097
```

RESULT 14

```
US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
```

```
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8439
TELEFAX: 301-309-8504
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1
Alignment Scores:
Pred. No.: 1.21e-44 Length: 1830121
Score: 475.50 Matches: 120
Percent Similarity: 50.74% Conservative: 52
Best Local Similarity: 35.40% Mismatches: 144
Query Match: 22.38% Indels: 23
DB: Gaps: 7
US-10-649-273-2 (1-414) x US-09-557-884-1 (1-1830121)
QY 39 ValLeuGlyIleGluThrSerCysAspAspThrAlaAlaValAlaSerGlyThrGly 58
Db 553208 ATCTTAGGCAATTTGAACCTTCTGATGAACCGCGCTGGATTTAGTGAAGAAAA 553149
QY 59 AsnValLeuGlyAlaIleHisSerGlnThrGluValHisLeuYsThrGlyGlyIle 78
Db 553148 GGAATTAATGCTAATCACTTATACATAAATTCGCTGACGACGATTAATGATGATG 553089
QY 79 ValProProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGluAla 98
Db 553088 GTTCTGAATTTAGCATCAGCGATCATATTCGAAAAACAGACCGCTTATTAAGCGCA 553029
QY 99 LeuSerAlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrIleLysPro 118
Db 553028 TTAGAGGAGCCCAATTTAACCGCAGCGATTTGATGATGATGCTTATACGAGTGGCCCT 552969
QY 119 GlYLeuAlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeu 138
Db 552968 GGGCTTGTGGCGCATTTCTTGTGGTCTGACGATGACATTCCTTAGCGCTTAGCTTGG 552909
QY 139 LysLysProPheIleProIleHisHisMetGluAlaHisAlaLeuThrIleArgLeu--- 157
Db 552908 AATGCTTCTGCGATTTGCTGATCATATGAAAGGCAATTTACTTGGCCCAATGCTTAT 552849
QY 158 ThrAsnLysValGluPheProPheLeuValLeuLeuIleSerGlyGlyHisCysLeuLeu 177
Db 552848 GACAAATTCACCGCACTTCTCTTTGTTGCTCTGTTGATGCGGTGGCCACATCAATTA 552789
QY 178 AlaLeuValGlnGlyValSerAspPheLeuLeuLeuGlyLysSerLeuAspIleAlaPro 197
Db 552788 GTGCGTGTGATGATGATGAAAAATGAAATGATGATGATGATGATGATGATGATG 552729
QY 198 GlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysProGlyCysSer 217
```

```

Db 552728 GGGCAGAGCTTGTGATTAATAACAGCAATAATTAAGTGAAGCTA---GATTATCA----- 552681
Oy 218 ThrMetSerGlyGlyValAlaIleGluHisIleValAlaGlyGlnGlyAsnArgPheHisPhe 237
Db 552680 -----GGTGGCGCGCGCTTCTCTCGTTTAGCGGAAAAAGATACGCCAAATCGTTTC 552630
Oy 238 AspIleIysProPheLeuHisHisAlaIysAsnCyAspPheSerPheThrGlyLeuGln 257
Db 552629 ACATTTCACGTCGATATGACAGATGTCGAGCGCTTGATTTTGAATTTCTTGTGTTTAAA 552570
Oy 258 HisValThrAspIleIleIleMetIysGlyIysGluGluGly-----IleGlyIys 275
Db 552569 ACATTTCGCGCAATATACAGTTATATCAAGCAATTAATAAACAGAGCGCAACTGATAGACAA 552510
Oy 276 GlyIleIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisIleThrMetAlaCyS 295
Db 552509 -----ACTTAACACAGATATTTGCTTATGCTTTCCAAAGATGCGGTGGAT 552465
Oy 296 HisIleValIysArgThrHisArgAlaIleLeuPheCysIysGlnArgAspLeuLeuPro 315
Db 552464 ACTCTTGCC-----ATTAATGTAG--CGTGCAATTGAAAAGAA 552425
Oy 316 GlnAsnAsnAlaValIleValAlaSerGlyGlyValAlaSerAsnPheIleArgArg 335
Db 552428 ACAGCGCTAATAACGTTTACTGATTTGGCGAGGGGTAGCGCAATATAAAAACTCCGAGAA 552368
Oy 336 AlaIeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCySPROPORProArgLeu 355
Db 552368 ACGCTTGGCGCACTTATATGCATAATTAGGTGGCGAAGTGTATTAATCCTCAACCTCAATT 552309
Oy 356 CysThrAspAsnArgIleIleMetIleAlaIleAsnGlyIleGluArgLeuArgAlaGly 374
Db 552308 TGTACAGATTAATGTGGCGATGATGCTTACACAGGTTTTTTACGTTTAAAAACAAGGT 552252

RESULT 15
U 09-643-990A-1/c
S Sequence 1, Application US/09643990A
P Patent No. 6528289
G GENERAL INFORMATION:
A APPLICANT: Robert D. Fleischmann
M Mark D. Adams
O Owen White
H Hamilton O. Smith
T TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof
N NUMBER OF SEQUENCES: 1
C CORRESPONDENCE ADDRESS:
A ADDRESSEE: Human Genome Sciences, Inc.
S STREET: 9410 Key West Avenue
C CITY: Rockville,
S STATE: MD
C COUNTRY: USA
Z ZIP: 20850
C COMPUTER READABLE FORM:
M MEDIUM TYPE: 3 1/2 inch diskette
C COMPUTER: Dell Pentium
O OPERATING SYSTEM: MS DOS V6.22
S SOFTWARE: ASCII Text
C CURRENT APPLICATION DATA:
A APPLICATION NUMBER: US/09/643,990A
F FILING DATE: 23-Aug-2000
C CLASSIFICATION: <Unknown>
P PRIOR APPLICATION DATA:
A APPLICATION NUMBER: 08/487,429
F FILING DATE: 1995-06-07
A APPLICATION NUMBER: 08/426,787
F FILING DATE: 1995-04-21
A ATTORNEY/AGENT INFORMATION:
N NAME: Kenley K. Hoover
R REGISTRATION NUMBER: 40,302
P REFERENCE/DOCKET NUMBER: PB186P1C1

```

```

! TELECOMMUNICATION INFORMATION:
! TELEPHONE: 301-610-5790
! TELEFAX: 310-309-8439
! INFORMATION FOR SEQ ID NO: 1:
!   SEQUENCE CHARACTERISTICS:
!     LENGTH: 1830121 base pairs
!     TYPE: nucleic acid
!     STRANDEDNESS: double
!     TOPOLOGY: linear
!   SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Alignment Scores:
Pred. No.: 1,21e-44 Length: 1830121
Score: 475.50 Matches: 120
Percent Similarity: 50.74% Conservative: 52
Best Local Similarity: 35.40% Mismatches: 144
Query Match: 22.38% Indels: 23
DB: 4 Gaps: 7

US-10-649-273-2 (1-414) x US-09-643-990A-1 (1-1830121)

QY 39 ValLeuGjIleGluThrSerCysAspAspThrAlaAlaValValAspGluThrGly 58
Db 553208 ATCTTAGCATTTGAATCACTTCCTGATGAAGACGGCCGCGCATTTATATGAGAAAA 553149
QY 59 AsnValLeuGjIleGluAlaIleHisSerGlnThrGluValHisLeuLysThrGlyIle 76
Db 553148 GGATTAATGCTATACCACTTATATCTCAAAATGGCTTCGATGACAGATTATGATGGTGG 553089
QY 79 ValProProlAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGluAla 98
Db 553088 GTTCCGTGAATTAGCATTCACGATATATATTCCGAAAAACAGACCTCTTATTAAAGCGCA 553029
QY 99 LeuSerAlaSerGlyValSerProSerAspLeuSerAlaIleAlaThrThrIleLysPro 118
Db 553028 TTAGAGGAAGCAATTTAAACCGCAGCATATTGATGTATTGCTTATACGATGGCCCT 552969
QY 119 GlyLeuAlaLeuSerLeuGjIleValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeu 138
Db 552968 GGGCTTGTGGGGCAATGCTTGTCCGTGTACAGATTGACGCTTTTACCTGATCGCTTGG 552909
QY 139 LysLysProPheIleProIleHisHisMetGluAlaHisAlaLeuThrIleArgLeu--- 157
Db 552908 AATGTTCTCGCATTTGGTGTATCATATGAAAGGCGATTATCTTCCGCCAATGCTTAT 552849
QY 158 ThrAsnLysValGluPheProPheLeuValLeuLeuIleSerGlyGlyHisCysLeuLeu 177
Db 552848 GACAAATTACCGCACCTTCTTTGTGCTGTTGTGATCGGGTGGCACACTCAATTA 552789
QY 178 AlaLeuValGlnGlyValSerAspPheLeuLeuGjIleLysSerLeuAspIleAlaPro 197
Db 552788 GTGCGCTTCGAAAGGTGTAGGAATAATGAAGTAAGGAATCTATATATGATGCTGT 552729
QY 198 GlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGluCysSer 217
Db 552728 GGGCAAGCCTTTGATTAACACACAAATTACTTGACTA---GATTATAC----- 552681
QY 218 ThrMetSerGjIleGlyAlaIleGlnHisLeuAlaLysGlnGlyAsnArgPheHisPhe 237
Db 552680 -----GGTGGCGCGCGCACCTTCTGCTTTAGCGGAAAAAAGTACGGCAATGCTTTC 552620
QY 238 AspIleLysProProLeuHisIleAlaLysAsnCysAspPheSerPheThrGlyLeuGln 257
Db 552629 ACATTTCCACGCTCCATGACACATGCTGCAGCGCTTGATTTTATTTCTCGTTTAAAA 552570
QY 258 HisValThrAspLysIleIleMetLysLysGjIleGlnGly-----IleGjIleLys 275
Db 552569 ACATTTCCGCAAAATACAGTTATACAGCAATTAACAGAGGAGGACCACTGATAGAGCA 552510
QY 276 GlyGlnIleLeuSerSerAlaAlaAspIleAlaIleThrValGlnHisIleThrMetAlaCys 295
Db 552509 -----ACTAAAGCATATTTGCTTATGCTTTCACAGATGCGGTGCTGAT 552465

```

```
Oy      296 HisLeuValIysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 315
      |||
      552464 ACTCTTGCC-----ATTAAATGTAG---CGTCATTGAAAGAA 552429
Oy      316 GlnAsnAsnAlaValIleuValAlaSerGlyGlyValAlaSerAsnPheTyrIleArgArg 335
      |||
      552428 ACAGGCTATAAACGTTTAGTGAATTCGGGAGGGGTGAGCGCAATATAAAAACTCCGAGAA 552369
Oy      336 AlaLeuGluIleLeuThrAsnAlaThrGlnCysThrIleuLeuCysProProArgLeu 355
      |||
      552368 AGCCTTGCGCAGCTTAATGCAAAATTTAGGTGCGCAAGTGTTTATCTCAACCTCAATTT 552309
Oy      356 CysThrAspAsnGlyIleMetIleAlaTrpAsnGlyIleGluArgLeuArgAlaGly 374
      |||
      552308 TGTACAGATPATGTGTGCGATGATTGCTTACACAGGTTTTTACGTTTAAAAACAAGT 552252
Db
```

Search completed: June 17, 2005, 02:00:31
Job time : 1592.65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 16, 2005, 17:01:19 ; Search time 3547.8 Seconds
(without alignments)
4441.797 Million cell updates/sec

Title: US-10-649-273-2

Perfect score: 2125
Sequence: 1 MLITFTAGVFPKSPKRYV.....DISKEVGEASIKVPOLKMEI 414

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+.p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10649273/runat.15062005.111416.6043/app.query.fasta.1.1429
-DB=EST -QFMT=fasta -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORH=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10649273.@CGN.1.1.8076.@runat.15062005.111416.6043 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:.*
1: gb_est1:.*
2: gb_est2:.*
3: gb_hic:.*
4: gb_est3:.*
5: gb_est4:.*
6: gb_est5:.*
7: gb_est6:.*
8: gb_gss1:.*
9: gb_gss2:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1827	86.0	2284	3	AK045669	AK045669 Mus muscu
2	1819	85.6	1622	3	AK011265	AK011265 Mus muscu
3	1634	76.9	1609	3	BC030671	BC030671 Mus muscu
4	1447	68.1	870	5	BQ423651	BQ423651 AGENCOURT
5	1353.5	63.7	852	5	BX391919	BX391919 AGENCOURT
6	1279	50.2	522	5	BC056102	BC056102 AGENCOURT
7	1173	55.2	701	2	BE740611	BE740611 601595739
8	1158	54.5	749	2	CK982692	CK982692 4115331 B
9	1144	53.8	800	7	CO738006	CO738006 SLIH03c19

10	1132	53.3	1109	5	BM907988	BM907988 AGENCOURT
11	1111	52.3	765	5	CF114247	CF114247 Shultzomi
12	1097	51.6	640	5	BQ636028	BQ636028 hd03d11.Y
13	1045	49.2	730	7	CN823245	CN823245 Oa.sp1bn
14	1038.5	48.9	661	7	CK833139	CK833139 4057170 B
15	1038.5	47.9	879	5	BQ433135	BQ433135 AGENCOURT
16	1017	47.9	1173	6	CD508917	CD508917 CDA93-B05
17	997	46.9	723	7	CK942291	CK942291 4065791 B
18	978	46.0	579	7	CK819035	CK819035 1f06d05.Y
19	966	45.5	822	7	CF257246	CF257246 pha008.Y
20	963	45.3	577	7	CV028547	CV028547 7089 Full1
21	954.5	44.9	808	7	CK792978	CK792978 AGENCOURT
22	954	44.9	833	5	BU159877	BU159877 AGENCOURT
23	951.5	44.8	728	4	BG436856	BG436856 60248845
24	935.5	44.0	879	5	BU256052	BU256052 603746167
25	935	44.0	871	2	BF308169	BF308169 601887411
26	925	43.5	658	7	CF362328	CF362328 829596.MA
27	923	43.4	811	7	CN232675	CN232675 MLB051B07
28	920	43.3	859	5	BU292595	BU292595 603604785
29	916	43.1	831	7	CK455183	CK455183 915735.MA
30	915	43.1	812	5	BU246489	BU246489 603784202
31	912	42.9	878	7	CK456228	CK456228 916887.MA
32	911	42.9	661	7	CF787040	CF787040 854903.MA
33	907	42.7	717	7	CK459683	CK459683 929895.MA
34	906	42.6	597	6	CB272391	CB272391 mai57g12.
35	902	42.4	533	4	BM126453	BM126453 1f06d05.Y
36	893	42.0	574	5	BM130908	BM130908 1f06d05.Y
37	893	42.0	687	7	CF360919	CF360919 827217.MA
38	879	41.4	737	7	CV073257	CV073257 AGENCOURT
39	867	40.8	616	6	CB438468	CB438468 687077.MA
40	857.5	40.4	789	5	BU242187	BU242187 603781023
41	838	39.4	637	7	CK941819	CK941819 4065407.B
42	831	39.1	853	7	CR444993	CR444993 CR444993
43	825.5	38.6	806	4	BG333362	BG333362 602431206
44	819.5	38.6	792	5	BU403563	BU403563 604138456
45	819	38.5	668	5	BU285782	BU285782 604166893

ALIGNMENTS

RESULT 1
AK045669
LOCUS
DEFINITION
AK045669 2284 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male corpora quadrigenima cDNA, RIKEN
full-length enriched library, clone B230219017 product: similar to
PUTATIVE STALOGICOPROTEIN TYPE 2 [Homo sapiens], full insert
Sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK045669
AK045669.1 GI:26337528
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
2
10349636
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
3
11042159
Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Atzawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitanai, T., Tashiro, H., Itoh, M.,
Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

TITLE	Pujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matubiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
JOURNAL	RIKEN Integrated sequence analysis (RISA) system--384-format
MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)
PUBMED	20530913
REFERENCE	11076661
REFERENCE	4
TITLE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
JOURNAL	Functional annotation of a full-length mouse cDNA collection
REFERENCE	Nature 409, 685-690 (2001)
REFERENCE	5
TITLE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
JOURNAL	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
REFERENCE	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 2284)
TITLE	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hizomoto, K., Hirooka, T., Hirozane, T., Horii, F., Imochi, K., Iehi, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, R., Sakazume, N., Sano, H., Saesaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akehira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
JOURNAL	Direct Submission
REFERENCE	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	CNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
COMMENT	Please visit our web site for further details.
COMMENT	URL: http://genome.gsc.riken.jp/
COMMENT	URL: http://fantom.gsc.riken.jp/
COMMENT	Location/Qualifiers
FEATURES	1. 2284
SOURCE	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="PANTOM,DB:B230219017"
	/db_xref="taxon:10090"
	/clone="B230219017"
	/sex="male"
	/tissue_type="corpora quadrigemina"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="adult"
	221. 1465
	/note="unnamed protein product; putative
	similar to PUTATIVE SIALOGLYCOPROTEASE TYPE 2 (Homo
	sapiens) (SPTR Q9H4B0, evidence: FASTV, 80%ID, 100%length,
	match=1242)"
	/codon_start=1
	/protein_id="BAC32450.1"
	/db_xref="gi:26337529"
	/translation="MLMRLPRAGAIIPKPKSKYVGLRFSVHPRTSLCHLVIGIEI
	SCDPTGAAVDENGVNGEALHSGOIVLTKTGIVPVAOOLHENTRIORIEYTLISAG
	RTPSDLSAATITKIKGLALSLGGLSLSDLVNVPKCPPIPIHMEHATITLRTN
	VEFFVLLISGCHLALVGVSDFFILAGSLDIPADMDLPAKRRSLIKHPCST
	MSGKALIEQLAKGNRHFHFTINPQNQAKNDFSTGLQHTLTKLKEEGEKGK

		polyA_signal		OIISSADIAA ¹ AAV ² ATAC ³ AKRT ⁴ HAII ⁵ FC ⁶ K ⁷ KNLS ⁸ PN ⁹ ANAL ¹⁰ VSG ¹¹ VAS ¹² NLY ¹³ IR	
		polyA_site		KALIEV ¹⁴ ANQ ¹⁵ CTLC ¹⁶ CP ¹⁷ PR ¹⁸ CT ¹⁹ NG ²⁰ MI ²¹ AM ²² NG ²³ IER ²⁴ L ²⁵ RAG ²⁶ GLV ²⁷ AD ²⁸ VED ²⁹ I ³⁰ RYE ³¹ PK ³² CE ³³ LP	
		polyA_site		GVDS ³⁴ SR ³⁵ SV ³⁶ AA ³⁷ IR ³⁸ VP ³⁹ RL ⁴⁰ KAL ⁴¹ "	
		/note="putative"		2262..2267	
		/note="putative"		2284	
		/note="putative"		2284	
ALIGNMENT					
US-10-649-273-2 (1-414) x AK045669 (1-2284)					
ALIGNMENT SCORES:					
Pred. No.:		5,12e-194		Length: 2284	
Score:		1827.00		Matches: 352	
Percent Similarity:		90.58%		Conservative: 23	
Best Local Similarity:		85.02%		Mismatch: 39	
Query Match:		85.98%		Indels: 0	
DB:		3		Gaps: 0	
QY	1	MetLeuIleLeuThrIlySThrAlaGlyValPhe	helsProSerIysArgLysValIlyr	20	
DB	221	ATGCTAAAGTTAGAAAGACAGACAGAGAGCTATTC	CCAGCCCAAGAAAGTAAAGTTAT	280	
QY	21	GlupPheIeuArgSerPheAsnPhHisProGlyThr	LeuPheIeuHisIlysIlleValLeu	40	
DB	281	GGATTTTAAAGAAATTTAGTTCCTCATCCGAACTC	CTCTTGTGCTAAATCGTCCCTG	340	
QY	41	GlyIllegIuthrSerCyAspAspThrAlaAlaVal	AlaValAlaAspGluThrGlyAsnVal	60	
DB	341	GGAATTGAAACACAGCTGATGACACAGAGCGGCTG	TGCGTGAATGAAATCGGAATGTG	400	
QY	61	LeuGlyGluAlaIleHisSerGlnThrGluValHis	IleuLysIlyThrGlyGlyIlyValPro	80	
DB	401	CTGGGGAGAGACGCGACTCCCAACTCAGGTTATCT	GAAACAGGGGGAATGTTTCT	460	
QY	81	ProAlaIagInIleuHisArgGluAsnIlegInArg	IleValGlnGluAlaIleuSer	100	
DB	461	CAAGTAGCTCAACAACCTCACAGAAATAATCAACA	AGAAATAGTAAAGAAACCTTTCT	520	
QY	101	AlaSerGlyValSerProSerAspLeuSerAlaIle	AlaThrThrIleLysProGlyIleu	120	
DB	521	GCGCTAGAAATCACCCACAGCGATCTCAGCAATTC	GCAATTCACATCAAAACCGGCACTG	580	
QY	121	AlaLeuSerLeuGlyValGlyLeuSerPheSerLeu	GlnIleuValGlyGlnIleuLysIlys	140	
DB	581	GCCCTAAGCTTGGAGATTGGCTTATCCTTTAGCTT	ACAGTAAATCAGTTTAAAAAG	640	
QY	141	ProPheIleProIleHisIleMetGlnAlaHisAla	IleuThrIleArgLeuThrAsnLys	160	
DB	641	CCATTCAATCCCGATTCAATCAGATGAGGCTACAG	CACTGATATTATTAAGGCTCAACCAATAA	700	
QY	161	ValGluPheProPheLeuValIleuLeuIleSerGly	GlnHisCyLeuLeuAlaIleuVal	180	
DB	701	GTAGAATTTCTTTTTRAGTCTTTTGAATTTCTGCG	CGGTCACTGCGCTGTGGCATTAATC	760	
QY	181	GlnGlyValSerAspPheLeuLeuLeuGlyLysSer	LeuAspIleAlaIleProGlyAspMet	200	
DB	761	CAAGGTGTTTCCGATTTCTGCTCTCTTGGGAAGT	CTTTGGACATATGACACAGGCGACATG	820	
QY	201	LeuAspLysValAlaArgArgLeuSerIleuIleLys	HisProGluInCySerThrMetSer	220	
DB	821	CTTGAACAAGGTGGCAAGAAAGCTTTCTTAAATCA	AAACATCCAGAAATGTTCTTAAATGAGT	880	
QY	221	GlyGlyLysAlaIleGlnHisIleuAlaLysGlnGly	AsnArgPheHisPheAspIleLys	240	
DB	881	GGTGGAAAGCTATAGAACAGTGGCCAAAGCGAAAT	TAGATTTCAATTTTATCTATCAAT	940	
QY	241	ProProLeuHisHisAlaLysAsnCyAspPheSer	PheThrGlyLeuGlnHisValAlaThr	260	
DB	941	CAACTATGCAGAATGCTTAAATTTGCAATTTCTTT	CTTTCACGGGACTTCAACATATTAAT	1000	
QY	261	AspLysIleIleMetLysGlyGluLysGluGluGly	IleGlyLysGlyGlnIleLeuSer	280	

Db	Accession	Version	Keywords	Source	Organism	Reference	Authors
Db	1001	GATNAGCTAATACACACCAAGGAAAGAAAGAACAGCATTTGAGGAAGGGCAATCTGTCA	1068				
Qy	281	SerialAalaaPpIleAlaAlaThValGlnHisThmethylAcylEHisLeuValLysArg	300				
Db	1061	TCAGCTCGAGACATTTGCTGCTCGCGGTACAGCATCCACACAGCGGTGCACCTTGGCAAAAGA	1120				
Qy	301	ThHisArgAlaIleLeuPheCysLeuGlnArgAspLeuLeuProGlnAsnAspAlaVal	320				
Db	1121	ACACATGCGGCTATTCGTTTGTGGACAGGAAATTTGCTCTCCAGCTAACGCCATTA	1180				
Qy	321	LeuValAlaSerGlyGlyValAlaSerAsnPheTyrIleArgArgAlaLeuGluIleLeu	340				
Db	1181	TTAGTTGATTCGTGAGAGGTGTGGCAAGTACTCTTGACATCCGAAAGCATTTGGAATTTGC	1240				
Qy	341	ThrAsnAlaThrGlnCysFthLeuLeuCysProProPolarGluCysFthAspArgGly	360				
Db	1241	GCAATATGCAACGACAGTGCACGTTGTTGTCTCACCTCCACAGCTGTGCATGCACATATGC	1300				
Qy	361	IleMetIleAlaTPrpArgGlyIleGlnArgLeuArgAlaGlyLeuGluIleLeuHisAsp	380				
Db	1301	ATCATGATGTCATGATGAATGAAATGAAGATTAAGTCTGCTGCGGCGTTTACATAT	1360				
Qy	381	IleGluGlyIleArgTyrGluProLysCysProLeuGlyValAspIleSerLysGluVal	400				
Db	1361	CTAAGACGATCCGATATGAGACCAAAATGCTCTCTTGAGATGACATATCCAGACACTT	1420				
Qy	401	GlyGluAlaSerIleLeuValProGlnLeuLysMetGluIle	414				
Db	1421	GCAGAGAGCTCCATPAAAGTACCGGCAATTAATAATGCACTT	1462				
RESULT 2							
AKO11265							
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
MEDLINE							
PUBMED							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
MEDLINE							
PUBMED							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
MEDLINE							
PUBMED							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
MEDLINE							
PUBMED							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
MEDLINE							
PUBMED							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
MEDLINE							
PUBMED							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
MEDLINE							
PUBMED							
REFERENCE							
AUTHORS							
TITLE			</				

TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1622)
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arawaka, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hizmoto, K., Hirakawa, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izaawa, M., Kenkawa, T., Kori, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Onno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, K., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-ree@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedic Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGGAGGACGACGATCCAGACCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGGAGGATCTTCAGTATTAATTAATATGCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI, 3' end: SstI. Host: SOLR.
FEATURES
source Location/Qualifiers
 1..1622
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM,DB:2610001M19"
 /db_xref="taxon:10090"
 /clone="2610001M19"
 /tissue_type="whole body"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="10 days embryo"
 207..1451
 /note="unnamed protein product; putative similar to PUTATIVE SIALOGLYCOPROTEIN TYPE 2 [Homo sapiens] (SPTR|Q9H4B0, evidence: FASTV, 80%ID, 100%length, match=1242)"
 /codon_start=1
 /protein_id="BAB27506.1"
 /db_xref="GI:12847276"
 /translation="MLMLRTAGAIIPKPKSKYVGLRPSVHPRITLSCHKLVIQIETSCDPTGAADVDETNVIGALHSQYVALKGGIVPVPAQVLENTORIVLETLNKA RIPSDDSAIATTKIPGLSIGVGSFSLQVLOVNPQPKPFIPIHMEAHALITLTSK VEPFVILISGCHLALVQUSDPILKSLDIADQMDIVARSLITIKPEGST MSGGAKIEQLKQGNRHFPTINPQNAKCDPSFTGLQHTIDKLITHEKEBEI BKS QIISADADIAAVQHAATFLAKRTHAALIFCQKNAISPNNAVLVSGVASVLYIR KALIVANQPTQTLCPBPPLCTDMGJIMIMWNGIERLRLACLTCTDEVEDINVERKCP LGVDSRVEAAIKVRLKVAL"
 1605..1610
 /note="putative"
 1622

DNA Sequencing by: Institute for Systems Biology
<http://www.systembiology.org>
 contact: anand@systembiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 66 Row: e Column: 10
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21312463
 This clone has the following problem: frame shifted.

FEATURES

source

1. 1609
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:1226118"
 /issue_type="Thymus gland, mouse"
 /clone_lib="Soares_thymus_2NBM1"
 /lab_host="DH10B"
 /note="Vector: pT73-Pac"

ORIGIN

Alignment Scores:

Pred. No.: 1 64e-172 Length: 1609
 Score: 1634.00 Matches: 327
 Percent Similarity: 83.86% Conservative: 21
 Best Local Similarity: 78.80% Mismatches: 40
 Query Match: 76.89% Indels: 27
 DB: 3 Gaps: 2

US-10-649-273-2 (1-414) x BC030671 (1-1609)

QY 1 MetLeuIleuThrIleuThrAlaGlyValPhePheIysProSerIysArgIleValTyr 20
 DB 229 ATCTATATGTTAAGAAGACAGACAGAGCATTCCTCCAAAGCCCAAGAGATTAAT 288
 QY 21 GluPheLeuArgSerPheAenPheHisProGlyThrLeuPheLeuHisIleValLeu 40
 DB 289 GGAATTTAAGAAGATTAGTTCATCCCAAGACTCTCTGTCATTAAGTGTCTCG 348
 QY 41 GlyTlLeuTherCysAspAspThrAlaAlaValValaAspGluThrGlyAsnVal 60
 DB 349 GGAATTGAACCAAGCTGTGTATGACACAGAGCGCTGTGGATGAATCGGGAATGTG 408
 QY 61 LeuGlyGluAlaIleIleSerGlnThrGluValHisIleu-LysThrGlyGlyIleValPr 80
 DB 409 CTGGGGGAACACTGCACTCCCAACTCAGGTTCACTGAACT-----GGATTGTTCC 462
 QY 80 oProAlaAlaGlnGlnLeuHisArgGluAenIleGlnArgIleValGlnGluAlaLeuSe 100
 DB 463 TCAGTAGTCTCAACAATTCAAGAGAAATATTCAAGATATGAGAAAGAACTTTTC 522
 QY 100 rAlaSerGlyValSerProSerSerAspLeuSerAlaIleAlaThrThrIleIysProGlyLe 120
 DB 523 TGCCTGAGATACACCCCAAGCACTCTCAGCAATTGCAACTACATCAAAACCGGCACT 582
 QY 120 uAlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuVal 140
 DB 583 GGCCCTTAAGCTGGAGCTGTATCTTAACTTACCTTCACTGATTAATCAATTTAAAAA 642
 QY 140 sProPheIleProIleHisIleMetGluAlaHisAlaLeuThrIleArgLeuThrAsnIly 160
 DB 643 GCCATTTCATCCGATTATCAATGAGGCTCAGCCACTGACTATTAGGCTCACCATAA 702
 QY 160 sValGluPheProPheLeuValLeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuVa 180
 DB 703 AGTAGAATTCCTTTTAGTCTTTTATTTCTGGGCTACAGCTGCTGGCTTGGCTTGT 762
 QY 180 lGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMe 200

DB 763 CCAAGCTTTCCCATTTCCGCTCCTTGGGAAGCTTTGGACATAGACACAGGAGCAT 822
 QY 200 tLeuAspIysValAlaArgArgLeuSerLeuIleIysHisProGluCysSerThrMetSe 220
 DB 823 GCTTGACAAGGTGGCAAGAACATTTCTTATATCAATCATCCAGATATGTTCTACATAG 882
 QY 220 rGlyGlyValAlaIleGluHisLeuAlaIleGlnGlyAsnArgPheHisPheAspIleIly 240
 DB 883 TGGTGAAGAAAGCTATAGAACAGTTGGCCAAAGACGGAATATGATTCATTTACTATCA 942
 QY 240 sProPheLeuHisHisAlaIysAsnGlyAspPheSerPheThrGlyLeuGlnHisValTh 260
 DB 943 TCCACTATGCAAGATGCTAAGATGGATTTTCTTACAGGACCTTCAACATATTAC 1002
 QY 260 rAspIleIleIleMetCysIysGluIysGluGlyIleGluIysGlyGlnIleLeuSe 280
 DB 1003 TGATTAAGCTATTAACACAGAGAAAGAAAGAAAGAGCATTTGAGAGGGCAATCTGTTC 1062
 QY 280 rSerAlaAlaAspIleAlaIleAlaThrValGlnHisThrMetAlaCysHisLeuValIysAr 300
 DB 1063 ATCAGCTGCAGACATTTGCTGCTGGGTACAGCATGCAACAGCTGCCACTTGCAGAAAG 1122
 QY 300 gThrHisArgAlaIleLeuPheCysIysGlnArgAspLeuProGlnAsnAsnAlaIys 320
 DB 1123 AACACATCGCCCTATTCTGTTTGCAGACAGAAATATGCTCTCTCAAGCTTAACGCAGT 1182
 QY 320 lLeuValAlaSerGlyValAlaSerAsnPheThrIleArgArgAlaLeuGluIlele 340
 DB 1183 ATTAGTGTATCTTGAAGCTGTTCAGAGTACTTGTATCTCCAAAGCATTTGGAATGT 1242
 QY 340 uThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLeuCysThrAspAsnG 360
 DB 1243 CGCAATGCAACGAGCTGACAGTGTGTGTCTGTCACCTCCAAAGCTGTGACATGACATG 1302
 QY 360 yIleMetIleAlaThrPheGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAs 380
 DB 1303 CATCATGATTGCA-----TGATGCTCTTGAAGTACATTAACAGAGAGT 1315
 QY 380 pIleGluGlyIleArgTyrGluProIysCysProLeuGlyValAspIleSerIysGluVa 400
 DB 1316 -----TGATGCTCTTGAAGTACATTAACAGAGAGT 1350
 QY 400 lGlyGluAlaSerIleIysValProGlnLeuIleMetGluIle 414
 DB 1351 TGCAGAGCTGCCATTAAGTACCCGATTAATAATGGCACTT 1393

RESULT 4
 BQ423651 870 bp mRNA linear EST 23-MAY-2002
 LOCUS AGENCOURT_7790948 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6065828
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ423651
 VERSION BQ423651.1 GI:21118966
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 870)
 NIH-MGC <http://imgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/BTP
 cDNA library Preparation: Life Technologies, Inc.
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LAM13342 row: 1 column: 21
 High quality sequence stop: 710.

FEATURES source

Location/Qualifiers

1..870

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6065828"

/cissue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1; Not: Site: 2; Salt: Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 7,296-152 Length: 870

Score: 1447.00 Matches: 285

Percent Similarity: 98.62% Conservative: 1

Best Local Similarity: 98.28% Mismatches: 2

Query Match: 68.09% Indels: 2

DB: 5 Gaps: 0

US-10-649-273-2 (1-414) x BQ423651 (1-870)

111 AAlaIleAlaThrThrIleLeuProGlyLeuAlaLeuSerLeuGlyValGlyLeuSerPhe 130

1 GCATTGCAACTCCATTAACAGAGAGCTTAAAGCTTGGAGTGGCTTATCATTTT 60

131 SerLeuGlnLeuValGlyLeuValLeuPheProIleLeuPheIleProIleHis 150

61 AGCTTACAGCTGTAGAGACAGTTAAAGCCATTCATTCATTCATCATATGAGGCT 120

151 HisAlaLeuThrThrIleLeuThrAsnLeuValGlyLeuPheProIleLeuLeu 170

121 CAGCACTTAATATGTTGGTTGACCAATAAAGTAGAATTTCTTTTAACTTTTAAAT 180

171 SerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGly 190

181 TCTGGAGGTACACTGTCTGTGGCATTTAGTTCAAGAGCTTCAAGTTTCTGCTTCTGGA 240

191 LysSerLeuAspIleAlaProGlyAspMetLeuAspValAlaArgArgLeuSerLeu 210

241 AAGTCTTTGGACATAGACAGAGTGAATCTTGACAAAGGTGGCAAGAGCTTTCTTTA 300

211 IleLeuHisProGlyCysSerThrMetSerGlyGlyValAlaIleGlnHisLeuAlaLys 230

301 ATAAACATCTCAAGAGCTCTCACCATGAGTGGGAAAGCCATTAAGAACTTTGGCCAAA 360

231 GlnGlyAsnArgPheHisPheAspIleLeuProIleLeuHisHisAlaLysAsnCysAsp 250

361 CAAGAAATGATTTCAATTTGACATCAACCTCTTGCAGACAGACATTCGCGCAAGTACAG 420

251 PheSerPheThrCysLeuGlnHisValThrAspLysIleIleMetCysGlyValGln 270

421 TTTTCTTTTACTGACTCAACGCTTACTGATTAATTAATTAAGAAAGAAAGAG 480

271 GlnGlyTyrIleGlnLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaThrValGln 290

481 GAAGGATTTGAGAAAGGCAAAATCCTGCTTGCAGACAGACATTCGCGCAAGTACAG 540

291 HisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPheCysValGln 310

541 CACACAAATGCAATGCTCTTGTGAAGAAAGACATCGGGCTATTCGTTTGTAAAGAG 600

311 ArgAspLeuLeuProGlnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsn 330

601 AGGAGACTTGTACTCAAAATATGCAATGCTGTTGATCTGTTGCTGCGCAAGTAC 660

331 PheTyrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCys 350

661 TTCTATATCGCAGAGCTCGAAATTTTAACAAACCAACACAGTGCACATTTGTGTGTGT 720

Qy 351 ProProProArgLeuCysThrAspAsnGlyIleMetIleAlaIleArgGlyIleGlnArg 370

Db 721 CCTCTCTCCAGACTATCATCATGATTAATGCAATTAATGATTAATGATTAATGAAAGA 780

Qy 371 LeuArgAlaGlyLeuGlyIleLeuHisAspIleGlnGlyIleArgGlyIleProLysCys 390

Db 781 CTAGCTCTGCTGCTGGCATTTTATCATGACATTAAGAGGACATCCCTATTAACCAATATG 840

Qy 390 sPro-LeuGlyValAspIleSerLys 398

Db 841 TCTCTCTGAGTATGACATTAACAAA 866

RESULT 5

BX391919 852 bp mRNA linear EST 28-APR-2004

LOCUS BX391919 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens

DEFINITION CDNA clone CS0DK001YE02 3-PRIME, mRNA sequence.

ACCESSION BX391919

VERSION BX391919.2 GI:46846154

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT On May 13, 2003 this sequence version replaced gi:30611736.

COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gasston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and Bclor V sites of the pCMVSPORT 6 vector. Library was normalized. Library was normalized.

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0BA1035ZEO7_CS03317_1&c=1240.r

FEATURES source

Location/Qualifiers

1..852

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DK001YE02"

/cell_type="HELA CELLS COT 25-NORMALIZED"

/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and Bclor V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:

Pred. No.: 2,416-141 Length: 852

Score: 1353.50 Matches: 274

Percent Similarity: 96.14% Conservative: 0

Best Local Similarity: 96.14% Mismatches: 8

Query Match: 63.69% Indels: 4

DB: 5 Gaps: 2

US-10-649-273-2 (1-414) x BX391919 (1-852)

Qy 72 HisLeuLys---ThrGlyGlyIleValProProAlaIleGlnGlnLeuHisArgGlnAsn 90

Db 850 CATTAAACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 798

Qy 91 IleGlnArgIleValGlnGlnAlaLeuSerAlaSerGlyValSerProSerAspLeuSer 110

```

Db      797 ATTCACGATAGTACAGAAAGCTCTTTCGACAGTGNAGTCTCTCCAAAGTACCTCTCA 738
Qy      111 AAlaIeAthrThrIleYsProGlyLeuAlaIeUSeSerLeuGlyValGlyLeuSerPhe 130
Db      737 GCATTTGCANCTACCACTAAACCCAGNAAGTCTTTTAAGCTGGAGAGGGCTTATCATTTT 678
Qy      131 SerLeuGlnLeuValGlyLeuValSerProPheIleProIleHisIsmetGluAla 150
Db      677 AGCTTACAGCTGGTAGAGAGAGTAAAGCCATTCATTCCTCCATTCATCATATGAGAGCT 618
Qy      151 HsAlaIeAthrThrIleArgLeuThrAsnLeuValGlnPheProPheLeuValLeuIle 170
Db      617 CATGCACTTACTATTAATTTGTTGACCAATTAAGTAGAATTTCTTTTATTTAGTTCTTTGATT 558
Qy      171 SerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGly 190
Db      557 TCTGGAGGCACTGCTGCTGTTGGCATTTAGTTCAAGAGTTTCAGATTTTCTGCTTCTTGA 498
Qy      191 LysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeu 210
Db      497 AGCTCTTTGACATACACACAGAGTGAATCTCTTGACAGAGTGGCAAGAGACTTTCTTTTA 438
Qy      211 IleValHisProGlyCysSerThrMetSerGlyGlyValAlaIleGluHisLeuAlaLys 230
Db      437 ATTAACCATCTCAGAGTGTCTCACCATGATGTGTGGGAAGCCATAGAACATTTGGCCAAA 378
Qy      231 GlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLysAsnCysAsp 250
Db      377 CAAAGAAATAGATTTCAATTTGACATCAACCTCCCTTGACATCAATGCTTAAATAATTTGAT 318
Qy      251 PheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysValGlyLeuGlu 270
Db      317 TTTCTTTTACTGACTTCAACAGCTTACTGATTAATTAATTAAGAAAAGAAAAGAG 258
Qy      271 GlnGlyIleGluLysGlyGlnIleLeuSerSerAlaIleAspIleAlaIleAthrValGln 290
Db      257 GAAGGATTTGAGAAAGGGCAAAATCTCTGCTTCAAGACAGACATTTGCTGCACAGTACAG 198
Qy      291 HisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGln 310
Db      197 CACACATGCGCATGCTCATCTTGTGMAAAGAACACATCGGGCTATTCGTGTTTGTAGACAG 138
Qy      311 ArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyValAlaSerAsn 330
Db      137 AGGAGCTTTGTACTCTCAAAATTAATGACATGCTGTGATCTGAGTGGTGGCAGAGTAC 78
Qy      331 PheArgIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCys 350
Db      77 TTCTAATTCGCGAGAGCTCTGGAATTTTAACCAACCAACAGTGCACCTTTGTGTGT 18
Qy      351 ProProProArgLeu 355
Db      17 CTCCTCTCCAGACTA 3

```

RESULT 6
 B0961028 LOCUS 922 bp mRNA linear EST 21-AUG-2002
 DEFINITION AGENCOURT 8863711 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6423902
 5' mRNA sequence.
 ACCESSION B0961028
 VERSION B0961028.1 GI:22376506
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NIH-MGC http://mhc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: DCTD/DTP/Gazdar

CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
 Plate: L10K2603 row: d column: 15
 High quality sequence stop: 584.
 Location/Qualifiers
 1..922
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6423902"
 /issue_type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 18"
 /note="Organ: lung; Vector: pOB7; Site 1: XhoI; Site 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN
 Alignment Scores:
 Pred. No.: 6.75e-133 Length: 922
 Score: 1279.00 Matches: 275
 Percent Similarity: 91.80% Conservative: 5
 Best Local Similarity: 90.16% Mismatches: 16
 Query Match: 60.19% Indels: 10
 Db: 5 Gaps: 3

US-10-649-273-2 (1-414) x B0961028 (1-922)
 Qy 80 ProProAlaAlaGlnGlnLeuHisArgGluAsnIleGlnArgIleValGlnGlnAlaLeu 99
 Db 3 CCTCCAGAGCTCAACAGCTTCAACAGAGAAATATTCACAGATGACAGAGCTCTT 62
 Qy 100 SerAlaSerGlyValSerProSerAspLeuSerAlaIleAthrThrIleLysProGly 119
 Db 63 TCTGCAAGTGGAGTCTCTCAAGTACCTCTCAGCAATGCAACATCAATTAACAGAGA 122
 Qy 120 LeuAlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuValGlyGlnLeuLys 139
 Db 123 CTTCCTTTAAGCTCTGGAGTGGCTTATCATTTAGCTTACACTGTGTAGACAGTTAAA 182
 Qy 140 LysProPheIleProIleHisIsmetGluAlaHisAlaLeuThrIleArgLeuThrAsn 159
 Db 183 AAGCATTCATTCATTCATCATATGAGAGGCTCATGCACTTACTATTAAGTTGACCAAT 242
 Qy 160 LysValGlnPheProPheLeuValLeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeu 179
 Db 243 AAAGTAGAATTTCTTTTACTTTTACTTTTCTTTGATTTCTGAGAGTACAGTCTTGGCATTA 302
 Qy 180 ValGlnGlyValSerAspPheLeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAsp 199
 Db 303 GTTCAAGAGGTTTCAATTTCTGCTTCTTGAAGTCTTTGGACATACACAGAGTAC 362
 Qy 200 MetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlyCys-SerThrMet 219
 Db 363 ATGCTTGAACAGAGTGGCAAGAGACTTTCTTTAATTAACATTCACAGATGACATCCACAT 422
 Qy 219 tSerGlyGlyValAlaIleGluHisLeuAlaLysGlnGlyValAsnArgPheHisPheAspIle 239
 Db 423 GAGTGTGGAAAGCCATAGCAATTTTGGCCAAACAGAAATAGATTTTCAATTTTACAT 482
 Qy 239 eLysProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGlyLeuGlnHisVal 259
 Db 483 CAAACTCTCCCTTGACATCATGCTAAATTTGTATTTTCTTTTACTGAGACTTCAACAGCG 542

TITLE G.P., Bosak, S., Rubenfield, M. and Gasbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut
JOURNAL Unpublished (2004)
COMMENT Contact: Ted S. Sonstegard
Bovine Functional Genomics Laboratory
Animal and Natural Resources Institute
Bldg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA
Tel: 3015048416
Fax: 3015048414

Email: tadesan@barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
0.000925 using options -trim all -trim fasta. Vector identified
by cross match using options -mismatch 12 -mismatch 18
Plate: 45 row: M column: 02
Seq primer: CCCAGTCACGACCTGTGTAACG
High quality sequence stop: 749.

FEATURES
source

1. 749
/organism="Bos taurus"
/mol_type="mRNA"
/strain="Holstein"
/db_xref="taxon:9913"
/clone="9BOV45_M02"
/sex="Male"
/tissue_type="Pooled"
/dev_stage="Multiple"
/lab_host="DH10B T1 phage resistant"
/clone_lib="BARC 9BOV"
/note="Organ: Abomasum; Vector: pagen-1; Site 1: EcorV;
Site 2: NotI; Equimolar amounts of mRNA extracted from
fundic and pyloric abomasums of 18 and 21 week old steers.
Exposure to Osteragia osteragi was initiated at 15 weeks
of age. fundic and pyloric abomasum"

ORIGIN

Alignment Scores:

pred. No.: 2.09e-119 Length: 749
Score: 1158.00 Matches: 225
Percent Similarity: 95.18% Conservative: 12
Best Local Similarity: 90.36% Mismatches: 12
Query Match: 54.49% Indels: 0
DB: 7 Gaps: 0

US-10-649-273-2 (1-414) x CK982692 (1-749)

QY 49 ThrAlaAlaValValAspGluThrGlyAsnValLeuGlyGluAlaLeuHisSerGln 68
DB 1 ACAGCAGCTCTGTAGTGAAGCGGAAATGTTTGGAGAGCAATCATTCCTCAA 60
QY 69 ThrGluValHisLeuLysThrGlyGlyTLeuValProAlaAlaGlnGlnLeuHisArg 88
DB 61 ACTGAAGTCATTTAAACAGGCGGAGATTAATCTCTCAGTAGCTCAACAGCTTCATGA 120
QY 89 GluAsnLLeuGlnArgLLeuGlnGlnAlaLeuSerLysSerGlyValSerProSerasp 108
DB 121 GAAATTAATTCACCCATAGTACAAAGAGCTCTCTCTCCAGTAGAAGTCTTCCAAAGTAA 180
QY 109 LeuSerAlaLeuAlaThrThrLLeuLysProGlyLeuAlaLeuSerLeuGlyValGlyLeu 128
DB 181 CTCCTACAGGTGCACACCATTAAGCCAGAGCTTGTCTTAAGCTTGCGGTAGTTTA 240
QY 129 SerPheSerLeuGlnLeuValGlyGlnLeuLysLysProPheLeuProLeuHisHisMet 148
DB 241 TCGTTTAGCTTACACATCGTAGACCAAGTTTAAAGCCCTTCATTCATTCATTCACATG 300
QY 149 GluAlaHisAlaLeuThrLLeuArgLeuThrAsnLysValGluPheProPheLeuValLeu 168
DB 301 GAGGCTCATGCACTCATATTAGGTAACAATTAAGTAGTAATTTCCGTTTATGTTCTT 360
QY 169 LeuLLeuSerGlyGlyHisGlyLeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeu 188
DB 361 TTGATTTCTGAGAGGTCAATGCTTTTGGCATTAAGTAGAGATTTCGATTTCTTCTT 420

QY 189 LeuGlyLysSerLeuAspLLeuAlaProGlyAspMetLeuAspLysValAlaArgArgLeu 208
DB 421 CTGGGAAGTCTTTGGACATAGCACAGGTACATGCTTGCAAGTAGAAGAGACTT 480
QY 209 SerLeuLLeuHisLysProGluLysSerThrMetSerGlyGlyValAlaLeuHisLeu 228
DB 481 TCTTTAATTAACATCCAGAGTGTCTCCACATGAGTGGGGAAGGCTATAGAACATTTG 540
QY 229 AAlaGlnGlnGlyAAsnArgPheHisPheAspLLeuLysProPheLLeuHisAlaLysAsn 248
DB 541 GCCAAACAGGAATATAGATTCATTTTCATTTCCAGCTCCCAAGCAACGCTTAAAT 600
QY 249 CysAspPheSerPheThrGlyLeuGlnHisValThrAspLysLLeuLLeuMetLysGlyGlu 268
DB 601 TGTATTTTCTTTTCTTTCGACTTCAACACGTTATGTAAGATGATATGATCANAGGAA 660
QY 269 LysGluGluGlyLLeuGluLysGlyGlnLLeuSerSerAlaAlaAspLLeuAlaThr 288
DB 661 AAAGAGGAAGGTATTCGAGCAGGAGGAGGAGGCTGTCTTCACTGAGACATTCGCTGGC 720
QY 289 ValGlnHisThrMetAlaCysHisLeu 297
DB 721 GTGACGACACCGGTGCTGCCACTT 747

RESULT 9
CO738006 800 bp mRNA linear EST 29-JUL-2004
LOCUS SLH03C19e02f1 squirrel heart library 1 Sperophilus lateralis cDNA
DEFINITION
clone 19e02 5', mRNA sequence.
ACCESSION
CO738006
VERSION
CO738006.1 GI:50825276
KEYWORDS
EST.

SOURCE

ORGANISM Sperophilus lateralis (golden-mantled ground squirrel)
Sperophilus lateralis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Scurinae;
Sperophilus.

1 (bases 1 to 800)

REFERENCE
AUTHORS Williams,D.R., Gracey,A.Y., Martin,S.L., Hughes,M.A., Li,W.,
Rogers,J. and Cossins,A.R.
TITLE Microarray analysis of transcriptional changes during hibernation
in the golden mantled ground squirrel, Sperophilus lateralis
JOURNAL Unpublished (2004)

COMMENT

Contact: Andrew R. Cossins
Laboratory for Environmental Gene Regulation
University of Liverpool
School of Biological Sciences, The Biosciences Building, Crown
Street, Liverpool, United Kingdom, L69 7ZB
Tel: +44(0)151-795-4510
Fax: +44(0)151-795-4431
Email: cossins@liv.ac.uk
Vector has been trimmed from this EST.
Plate: 19 row: e column: 02
Seq primer: pflc T7 (5'-AATACGACATCATAGG-3')
High quality sequence stop: 800.

FEATURES

source
1. 800
/organism="Sperophilus lateralis"
/mol_type="mRNA"
/db_xref="taxon:76772"
/clone="19e02"
/sex="Male & female"
/tissue_type="Heart"
/dev_stage="Adult"
/lab_host="E.coli Electromax DH10B"
/clone_lib="squirrel heart library 1"
/note="Vector: pflc; Site 1: SalI GTGACG; Site 2: BamHI
GGATCC; Normalized and subtracted cDNA library prepared
from heart of hibernating and summer animals"

ORIGIN

Alignment Scores:

Pred. No.: 8,74e-118 Length: 800
 Score: 1144.00 Matches: 226
 Percent Similarity: 97.05% Conservative: 4
 Best Local Similarity: 95.36% Mismatches: 7
 Query Match: 53.84% Indels: 0
 DB: 7 Gaps: 0

US-10-649-273-2 (1-414) x COJ38006 (1-800)

QY 1 MetLeu1leuLeuThryThralaGlyValaPhePhePheProSerLySaArgLyValTyr 20
 DB 88 ATGCTTAATTTGAATTAAGACAGCCGAGTTTCTTTTAAACATCCAAAGAAAGTTGAT 147
 QY 21 GluPheLeuArgSerPheAenPheHisProGlyThrLeuPheLeuHisLyValLeu 40
 DB 148 GCATTTTAAAGAGTTTAAATTTTCACTCTGGAACACTATTTCTTCAATAAATGATTTG 207
 QY 41 Gly1leGluThrSerCysAspAspThrAlaAlaValaValaPheGluThrGlyValaVal 60
 DB 208 GGTATTGAAACAGAGCTGTGATGATACAGCAGCTGCGGTGTAGATGAAGTGAAGTGTG 267
 QY 61 LeuGlyValAlaIleHisSerGlnThrGluValHisLeuLySerGlyValLeuValPro 80
 DB 268 TTGGGGAAGCAATACATTTCCAAACCTGAAGTTCAATTTAAACAGGTGGGATTAATCTT 327
 QY 81 ProAlaAlaGlnGlnLeuHisArgGluAsn1leGlnArg1leValaGlnGluAlaLeuSer 100
 DB 328 CCAGTAGCTCAACACTTCTATAGAGAAACATTCACAGCATATGACAGAGAGCTTTGT 387
 QY 101 AlaSerGlyValaSerProSerAspLeuSerAlaIleAlaThrThr1leLyProGlyLeu 120
 DB 388 GCCAGTGGAAATATCTCCAGTGAACCTCTCAGCATTCACATTAAGCCAGGACTT 447
 QY 121 AlaLeuSerLeuGlyValaGlyLeuSerPheSerLeuGlnLeuValaGlyGlnLeuLyVal 140
 DB 448 GCTTTAGCTTGAAGAGTAGGCTTATCATTTAGTTTACAGCTGTGAGACCGAGTGAAGAAAG 507
 QY 141 ProPhe1lePro1leHisHisMetGluAlaHisAlaLeuThr1leArgLeuThrAsnLyS 160
 DB 508 CCATTCATCCCATTCATCATCATATGAGAGCTCAGCGACTTCATTTAGTTGACCAATTA 567
 QY 161 ValGluPheProPheLeuValaLeuLeu1leSerGlyGlyHisCysLeuLeuAlaLeuVal 180
 DB 568 GTAGAGTTTCCATTTTATGATTTCTTTGATTTCTGAGAGTCACTGTTCTTGAGATTTGTT 627
 QY 181 GlnGlyValaSerAspPheLeuLeuLeuGlyLySerLeuAsp1leAlaPProGlyAspMet 200
 DB 628 CAAGAGTTTCAGATTTTCTGCTTTGGAAGCTTTTGACATYTGCACAGAGTGACATG 687
 QY 201 LeuAspLyValaAlaArgArgLeuSerLeu1leLyHisProGluCysSerThrMetSer 220
 DB 688 CTTGACAGAGTACAGAGAGACTTTCTTTAATCAACATCAAGATCTCCACCATATAGT 747
 QY 221 GlyGlyLyValaAlaIleGlnHisLeuAlaLyGlnGlyAsnArgPheHisPhe 237
 DB 748 GGCGGGAAGGCTATAGACATTTGGCCAAACAGAGAAACAGATTTTCAATTT 798

RESULT 10
 BM907988 1109 bp mRNA linear EST 12-MAR-2002
 LOCUS AGENCOURT 6707465 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5744534
 DEFINITION 5', mRNA sequence.
 ACCESSION BM907988
 VERSION BM907988.1 GI:19358367
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1109)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Straubeberg, Ph.D.
 Email: cgsdbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 plate: LLAM12767 row: b column: 07
 High quality sequence stop: 602.
 Location/Qualifiers

FEATURES

SOURCE

1. 1109
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5744934"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /clone_1lb="NIH_MGC_119"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 3.28e-116 Length: 1109
 Score: 1132.00 Matches: 241
 Percent Similarity: 88.89% Conservative: 15
 Best Local Similarity: 83.68% Mismatches: 23
 Query Match: 53.27% Indels: 9
 DB: 5 Gaps: 4

US-10-649-273-2 (1-414) x BM907988 (1-1109)

QY 1 MetLeu1leuLeuThryThralaGlyValaPhePhePheProSerLySaArgLyValTyr 20
 DB 129 ATGCTTAATTTGAATTAAGACAGCCGAGTTTCTTTTAAACATCCAAAGAAAGTTTAT 188
 QY 21 GluPheLeuArgSerPheAenPheHisProGlyThrLeuPheLeuHisLyValLeu 40
 DB 189 GAATTTTAAAGAGTTTAAATTTTCACTCTGGAACACTATTTCTTCAATAAATGATTTG 248
 QY 41 Gly1leGluThrSerCysAspAspThrAlaAlaValaValaPheGluThrGlyValaVal 60
 DB 249 GGAATTTAAACTAGTTATGATGATACAGCACTCTGTGTGAGTGAAGAACTGAAATGTG 308
 QY 61 LeuGlyValaAlaIleHisSerGlnThrGluValHisLeuLySerGlyValLeuValPro 80
 DB 309 TTGGGAAGACAAATACATTTCCAAACTGAAGTTCAATTAATAAACAAGGTGGATTTGCT 368
 QY 81 ProAlaAlaGlnGlnLeuHisArgGluAsn1leGlnArg1leValaGlnGluAlaLeuSer 100
 DB 369 CCAGCACTCAACAGCTTCAAGAGAAATATTAAGAAATGACAGAGCTCTTTCT 428
 QY 101 AlaSerGlyValaSerProSerAspLeuSerAlaIleAlaThrThr1leLyProGlyLeu 120
 DB 429 GCCAGTGAAGTCTCCAGAGTACCTTCAGCAATGTGCACTATCAATTAACCAAGACTT 488
 QY 121 AlaLeuSerLeuGlyValaGlyLeuSerPheSerLeuGlnLeuValaGlyGlnLeuLyVal 140
 DB 489 GCTTTAGCTTGGAGAGGCTTATCATTTAGCTTACAGCTGTGAGACAGCTTAAAAAG 548
 QY 141 ProPhe1lePro1leHisHisMetGluAlaHisAlaLeuThr1leArgLeuThrAsnLyS 160
 DB 549 CCATTCATCCCATTCATCATATGAGAGCTCAGCATCACTACTATTAGTTGACCAATTA 608

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
FEATURES
source

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 640)
Wistow, G., Bernstein, S.L., Wyatt, M.K., Ray, S., Behal, A., Touchman, J.W., Bouffard, G., Smith, D., and Peterson, K. Expressed sequence tag analysis of human retina for the NEURBank Project: Rebinding, an abundant, novel retinal cDNA and alternative splicing of other retina-preferred gene transcripts
Mol. Vis. 8 (4), 196-204 (2002)
22103461
12107411
Contact: Wistow, G.
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: gsteemehelix.nih.gov
Plate: 03 row: d column: 11
Seq primer: M13Rpl reverse primer (AB1).
Location/Qualifiers
1. 640
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hd03d11"
/issue_type="Retina"
/dev_stage="Adult"
/lab_host="EMD10B"
/clone_11b="Human Retina cDNA (Un-normalized, unamplified): hd/he"
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue was dissected from two 80 year old donors with no observed eye disease. 100ug of total RNA was used for library construction. A directionally cloned cDNA library in the pSPORT1 vector (Life Technologies) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the Superscript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor
[5'-PGACTGATGTTGATCGCGCGCCGCTT15-3']. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:
Pred. No.: 1.22e-112 Length: 640
Score: 1097.00 Matches: 209
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 51.62% Indels: 0
Gaps: 0

US-10-649-273-2 (1-414) x BQ636028 (1-640)

ORIGIN

206 ARGAGVLEUSeRleuilelyshisPRoGluCyseSerThnMetserGlylyValaile 225
|||
3 AAGAGCTTCTTTCTTAATAAACATCCAGATGCTCCACATGATGCTGGAAAGCATA 62
|||
226 GIUHLiLeuAlaLysGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 245
|||
63 GAACATTTGGCCAAACAGAAATAGATTTCATTTGACATCAAACTCCCTGCATCAT 122
|||
246 ALaLysAsnCyAspPheSerPheThnGlyLeuGInHisValThraSplylleMet 265
|||
123 GCTAAATAATTTGATTTCTTTTACTGCACCTTCAACAGCTTCTGTAATAATATATG 162
|||
266 LysLysGInLysGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 285
|||
183 AAAAAAGAAAAAGAGAGATTTGAGAAAGGGCAATCTGCTTCAAGACAGACAT 242
|||

ORIGIN

Alignment Scores:
Pred. No.: 1.08e-106 Length: 730
Score: 1045.00 Matches: 204
Percent Similarity: 93.51% Conservative: 12
Best Local Similarity: 88.31% Mismatches: 15
Query Match: 49.18% Indels: 0
Gaps: 0

US-10-649-273-2 (1-414) x CN823245 (1-730)

ORIGIN

286 ALAlaThrValGInHisThrMetAlaCyHisLeuValLysArgThrHisArgAlaile 305
|||
243 GCTGCCACAGTACAGCACAATGCGATCTCTTGTGAAGAACACATCGGGCTAAT 302
|||
306 LeuPheCysLysGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 325
|||
303 CTTGTTTGTAGACAGAGAGCTTTACTTCAAAATATGAGTACGTTGCAATCGGT 362
|||
326 GIVAlaLaseRanPheThnGlyLeuArgAlaLeuGInLysLeuThnAsnAlaThnGIn 345
|||
363 GGTCTCCAGATTAATCTTATATCCGACAGCTCTGGAATTTTAAACAGCAACAG 422
|||
346 CysThrLeuLeuCySPProProProArgLeuCySPThraSplylleMetIleAlaTrp 365
|||
423 TGCACTTTGTTGTTCTCTCCCTCCAGACTATGACATGATATGATATGATGATG 482
|||
366 AsnGInLysLeuArgAlaGInGInGInGInGInGInGInGInGInGInGInGInGIn 385
|||
483 AATGTAATTAAGAAACATACGCTGCTGGCATTTTACATGACATGAAAGCATCCGC 542
|||
386 TyrGInProLysCySPProLeuGInValAspIleSerLysGInValGInLysAlaSerIle 405
|||
543 TATGAACCAAAATATCTCTTGGAGTACATATCAAAAGAGTTGAGAACTTCCATA 602
|||
406 LysValProGInLeuLysMetGInLys 414
|||
603 AAGTACCAACATTAATAATGAGATA 629
|||

RESULT 13
CN823245 730 bp mRNA linear EST 02-JUN-2004
LOCUS
DEFINITION
Oa splbn_04N08 M13reverse Sheep spleen/brain pSport1 library Ovis aries cDNA clone Oa splbn_04N08 5', mRNA sequence.
CN823245
ACCESSION
CN823245.1 GI:47951314
VERSION
EST.
KEYWORDS
SOURCE
ORGANISM
Ovis aries (sheep)
Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Caprinae; Ovis.
1 (bases 1 to 730)
Gossner, A. and Hopkins, J.
Ovine spleen/brain cDNA library
Unpublished (2004)
JOURNAL
COMMENT
Contact: J Hopkins
Veterinary Biomedical Sciences
University of Edinburgh
Summerhall Square, Edinburgh, EH9 1QH.
Email: j.hopkins@ed.ac.uk
Plate: 04 row: N column: 08
Seq primer: M13reverse
High quality sequence start: 6
High quality sequence stop: 550.
Location/Qualifiers
1. 730
/organism="Ovis aries"
/mol_type="mRNA"
/db_xref="taxon:9940"
/clone="Oa splbn_04N08"
/clone_11b="Sheep spleen/brain pSport1 library"
/note="Vector: pSport1"

Qy	96	GlGlnIuIaLeuSerAlaSerGlyValSerProSerAspLeuSerAlaIleAlaThr	115
Db	37	CAAGAGCTCTCTCTCCAGTGAAGTCTCTCCAGTAACTCTGACAAATTCGCAACTACC	96
Qy	116	IleYsPProGlyLeuAlaLeuSerLeuGlyValGlyLeuSerPheSerLeuGlnLeuVal	135
		:::	
Db	97	GTAAGCGCAGGACCTTGCTTAAAGCTTGCGGTAGGTTATCATTTAGTCAACTGGTA	156
Qy	136	GlyGlnLeuLeuYsYsPProPheIleProIleHisIshwGluAlaHisIleuThrIle	155
Db	157	GACCAAGTTTAAAAAGCCCTTCATTCCTCCATTCATCACATGAGAGGCTCAACGACTTACTATT	216
Qy	156	ArgLeuThrAnuYsValGluPheProPheLeuValLeuLeuIleSerGlyHisIcys	175
Db	217	AGGTTAACAAATTAAGTAGTAATTCATTTTATGTTCTTTGATTTCTGAGAGTCATATG	276
Qy	176	LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyYsSerLeuAspIle	195
Db	277	CTTTGGCATTAGCTAGAGAGAGTTACAGATTTTCTTCTTTGGAAAGTCTTTGGACATA	336
Qy	196	AlaProGlyAspMetLeuAspYsValAlaArgArgLeuSerLeuIleYsHisProGlu	215
Db	337	GCACACAGGCCGACATGCTTGACAGAGTACCAAGACCTTGCTTATATAAACATCCAGAG	396
Qy	216	CysSerThrMetSerGlyGlyYsAlaIleGluHisLeuAlaYsGlnGlyAsnArgPhe	235
Db	397	TGCTCCACCATGAGAGGCGGAGGCTATTGAAACATTTGGCCAAACAGGGAATATGATTTG	456
Qy	236	HisPheAspIleYsProProLeuHisHisAlaYsAsnYsAspPheSerPheThrGly	255
Db	457	CATTGTGATTTCCCAACCTCCATGCACAGCTGCTAAATATGATTTTCTTTTCGGA	516
Qy	256	LeuGlnHisValThrAspYsIleIleMetYsYsGluYsGluGluGlyIleGluYs	275
Db	517	CTTCAACATGTTATTATGATTAAGTATTAATTCAGAGAGAAAAAGAGAGCATTCAGACAG	576
Qy	276	GlyGlnIleLeuSerSerAlaAlaAspIleAlaIleThrValGlnHisThrMetAlaCys	295
Db	577	GGCGAGATCCGTCTTCAAGTCGACACATTCGTCTGTCAGTCCACACACACGTAACCTGC	636
Qy	296	HisLeuValYsArgThrHisArgAlaIleLeuPheCysYsGlnArgAspLeuPro	315
Db	637	CACGTGCGAATAAGAACACATCGCCCTATTCTGTTCTGCAAGACAGAGAGCTTGTTACAT	696
Qy	316	GlnAsnAsnAlaValLeuValAlaSerGlyGly	326
Db	697	CAAGTAAACGACGACTGTTGTATCTGAGAGC	729
RESULT 14			
CR833139	661 bp	mRNA	linear
LOCUS	40571710	BARC_8BOV_Bos_taurus	cdna clone 8BOV_20020 5', mRNA
DEFINITION	Sequence.		
ACCESSION	CR833139		
VERSION	CR833139.1	GI:45061861	
KEYWORDS	EST.		
SOURCE	Bos taurus	(cow)	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
	Bovinae; Bos.		
	1 (bases 1 to 661)		
REFERENCE	Baumann,R.G., Baldwin,R.L., Sonstegard,T.S., Van Tassel,J.C.P. and		
AUTHORS	Matsumura,I.I.,K.		
TITLE	Construction and Analysis of a cDNA Library Generated From		
	Intestinal Muscle and Epithelial Tissues of Holstein Cattle		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: Richard G. Baumann		
	Bovine Functional Genomics lab		
	ANRI		
BUDG	162: BARC-EAST, Beltsville, MD 20705, USA		
	Tel: 3015048604		

Email: rbaumann@nrl.barc.usda.gov
Fax: 3015048744

Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim_alt ''-trim_faata. Vector identified by cross_match using options -mismatch 12 -minscore 18 Plate: 20 row: O column: 20 Seq primer: CCTATTAGTGCACCTATGAGAACC High quality sequence stop: 661.

Location/Qualifiers
1..661
 /organism="Bos taurus"
 /mol_type="mRNA"
 /strain="Holstein"
 /db_xref="taxon:9913"
 /cloned="8BOV_20020"
 /sex="Female"
 /tissue_type="Epithelial, Muscle"
 /dev_stage="Lactating, Neonatal"
 /lab_host="DH10B Tona"
 /cloned_jid="BARC 8BOV"
 /note="Organ: Intestine; Vector: pCMVSport6.1. Site_1:
NotI; Site_2: EcoRI; Normalized cow cDNA intestinal library in pCMVSport6.1, constructed from equimolar mRNA pools derived from 5 sources, 4 lactating intestinal, 1 neonatal intestinal 4/5 Lactating, Proximal Duodenum,
Jejunum, Distal Ileum, Colon, 1/5 Neonatal, Proximal Duodenum, Jejunum, Distal Ileum"

ORIGIN

Alignment Scores:

Pred. No.:	4,98e-106	Length:	661
Score:	1038.50	Matches:	203
Percent Similarity:	95.46%	Conservative:	8
Best Local Similarity:	91.86%	Mismatches:	9
Query Match:	48.87%	Indels:	1
DB:	7	Gaps:	1

US-10-649-273-2 (1-414) x CKR3J139 (1-661)

OY ProGIyThrleuPheLeuHsiLySileValleuglyllegluThrSerCyAspaSPthr 49
 |||||:::|||||||:::|||||
Db 2 CCCCC---AATTCCTCATTAACAATAAGTTGGGAATCGAAACCAAGTTGTATGATACA 58

OY Alaalaaavalvalaspgluthrghyaenvalleunglygluaallehisserginthr 69
 |||||
Db 59 GCAGTGCTGTAGTAGTGAAGACTGGAATGYTTTTGGGAGAACCAATACATCCCAAAT 118

OY GlUvaIHslseuLYethrgylgilyllevalpropoalaalaginInueuhlsarsglu 89
 |||||
Db 119 GAAGTTCATTAAAAACAGTGGGANtATTCCTCCAGTACTCACAGCTTCATTAGGAAA 178

OY AsnlieglnatgilvaelinglnualaueseralasergiliyalserproseraspLeu 109
 |||||
Db 179 AATATTCAACGCATGATGACAAABAGCTCTCTGCCAGAGAACTCTCCAATGAACTC 238

OY SerialaleiafthrThrrilleysProgiyleualaleuserteuuglylaiglyeuser 129
 |||||
Db 239 TCAGAGATTGCACACCACTAATAAGCCAGGACTTGAAGCTTTAGGCGCTAGGCTTAATCG 298

OY Phesertleuinleuvaiglynleuulsyrprophelleprojlehshsmetglu 149
 |||||
Db 299 TTtAccttACCAcCTGTGtagacccagtttaaaaaggccccttcattcccatccaatgcagag 358

OY AlahisalieuthrllekrgleuthrreaulysvaiglUpheProphelauvallleulu 165
 |||||
Db 359 GCTCATGCACCTACCTATTNGTttaaCaabaTaMGrtGaAttTCCgtTTtAgAtTCttTtg 418

OY Illesergilgyihicysleuleuallaleuvailnginliyvaliseraspheuleulen 189
 |||||
Db 419 ATTCTGgaGGtcaTtggcttTtTtggcataTgTTagAGGagtTtCaGAtttTtTtCTTCT 478

OY GliyvsertleuasptllleaiprogiYaapMetleuasplyVaAlaagaagleuer 209
 |||||

Db 479 GGGAGCTCTTTGGACATAGACCAAGGTGACATGCTTGACAGAGTACAGAAAGACTTTCC 538
Qy 210 LeuilelyshisProglucysSerthrmetserclyglysalailegluhsleuval 229
Db 539 TTTATTAATCATTCAGAGTCTCCACCATAGAGGGCGGAAGGTATAGAAACATTTGGCC 598
Qy 230 lvsGlnGlyAsnArgPhehisPheaspillelyserProleuhsishisalaValysAsnGys 249
Db 599 AAACAGGGAATATGATTCATTTTGATTTCCACCTCCATGCAACGTCTAAATAATGT 658
Qy 250 ASD 250
Db 659 GAT 661
RESULT 15
BQ433135 879 bp mRNA linear EST 24-MAY-2002
LOCUS AGENCOURT_7760756 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6016098
DEFINITION 5', mRNA sequence.
ACCESSION BQ433135
VERSION BQ433135.1 GI:21172211
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 879)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM13213 row: d column: 19
High quality sequence stop: 674.
Location/Qualifiers
1. 879
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6016098"
/issue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site: 1; Noti;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.: 1.39e-103 Length: 879
Score: 1018.50 Matches: 210
Percent Similarity: 94.67% Conservative: 3
Best Local Similarity: 93.33% Mismatches: 9
Query Match: 47.93% Indels: 3
DB: 5 Gaps: 1
US-10-649-273-2 (1-414) x BQ433135 (1-879)
Qy 1 MetLeuileleuthrlystrhalaGlyValPhePheIysProserlylsarglyvaltyr 20
Db 201 ATGCTAATCTTGAAGTCAAGAGGTTTATTAACATCAAAAGAAAGTTTAT 260
Qy 21 GluPheleuArgserPheAsnPhehisProGlyThrleuPheleuhslyslleValleu 40

Db 261 GAATTTTAAAGAACTTTAAATTTTCATCTCGAAGACATATTTCTTCAATAAATAGTATG 320
Qy 41 GlyilegluThrserCysAspAspThrAlaAlaValValAspGluThrGlyAsnVal 60
Db 321 GAATTAAGAACTAGTTGTGATGATAGACAGAGCTGTGGTGATGAACCTGAAAGTGT 380
Qy 61 leuGlyGluAlailehisserGlnThrGluValhisleuLysThrGlyIleValPro 80
Db 381 TTGGAGAAACCAATACATCTCCCAAGCTGAAGTTCAATTAATAAAGGAGGAGTGTCT 440
Qy 81 ProhlaIaGlnGlnleuhsisargGluAsnIleGlnArgIleValGlnGluAlaLeuSer 100
Db 441 CCAGCACTCAACAGCTTCACAGAAATATTCAGAAATAGTACAGAAAGCTTTCT 500
Qy 101 AlaserGlyValserProserAspSerAlaIleAlaThrThrIlelyProGlyLeu 120
Db 501 GCCAGTGAAGTCTTCCAGAGTACCTTCAGCAATTCACATACATCAAAACAGACTT 560
Qy 121 AlaleuSerleuGlyValGlyLeuSerPheSerleuGlnleuValGlyGlnleuLys 140
Db 561 GCTTAAGCTGGAGAGTGGCTTATCATTTAGCTTACAGCTGTAGACAGTAAAG 620
Qy 141 ProPheIleProIlehisshismetGluAlaHisAlaLeuThrIleArgleuThrAsnLys 160
Db 621 CCATTCATTCCTCCATTCATCATATGAGGCTCATGACCTTACTATTAGGTGACCAATTA 680
Qy 161 ValGluPheProPheleuValleuLeuIleSerGlyGlyhisCysleuLeuAlaVal 180
Db 681 GTAGAAATTTCTTTTATGTTCTTTGATTTCTGAGAGTCACTGTCTGTGGCATTA 740
Qy 181 GlnGlyValserAspPheleuLeuGlyLysSerleuAspIleAlaProGlyAspMet 200
Db 741 CAAGAGATTCAGATTTTCTGCTTCTTGAAGCTTTGGACATAGCACAGGTGACATG 800
Qy 201 -LeuAspLysValAlaArgArgleuSerleuIle---LysHis-ProGluCysSerthrm 219
Db 801 GCTTACAGAGGTGGCAAGAAAGACTTCTTGTATTAATAAATTCACAAAGTCTCCCC 860
Qy 219 etserGlyGly 222
Db 861 ATGGAANGGGT 871

Search completed: June 17, 2005, 01:29:52
Job time : 3564.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 16, 2005, 17:01:19 ; Search time 2288.07 Seconds
(without alignments)
4441.797 Million cell updates/sec

Title: US-10-649-273-2_COPY_148_414
Perfect score: 1385
Sequence: 1 MEAHATIRLTKVKEPFLV.....DISKEVGASIKVPOKMEI 267

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ .p2n.model -DEV=xlp
-Q/cgn2.1/USPTO.spool_p/US10649273/runat.15062005.11416.6043/app.query.fasta.1.1429
-DB=EST -OPMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10649273 @CGN 1.1 8076 @runat.15062005.11416.6043 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_g881: *
9: gb_g882: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1263	91.2	870	5	BQ423651	AGENCOURT
2	1224	88.4	2284	3	AK045669	Mus muscu
3	1216	87.8	1622	3	AK011265	Mus muscu
4	1097	79.2	640	5	BQ636028	hd03d11.y
5	1075	77.6	852	3	BX391919	BX391919
6	1071	77.3	1605	3	BC030671	Mus muscu
7	1053	76.0	701	2	BE740611	601595739
8	944	68.2	922	2	BQ961028	AGENCOURT
9	925	66.8	658	7	CF362328	829596 MA

10	916	66.1	822	7	CF257246	phs008.90
11	838	60.5	637	7	CF941819	4065407.B
12	830	59.9	597	6	CB272391	ma157912
13	819.5	59.2	792	5	BU403563	BU403563
14	816	58.9	730	7	CB823245	Ca.sp1dn
15	789	57.0	545	1	AV602901	AV602901
16	786	56.8	866	5	BU127463	603114407
17	784	56.6	490	6	CB852881	UI-CF-FNO
18	749	54.1	736	7	CK365185	CK365185
19	749	54.1	1173	6	CD508917	CD508917
20	746	53.9	723	5	BU261251	603502215
21	722	52.1	812	5	BU264689	603784202
22	717	51.8	634	2	AM601179	AM601179
23	709	51.2	484	1	AJ670918	AJ670918
24	708	51.1	736	6	CA057753	ssal19D54
25	707.5	51.1	701	5	BU621780	UI-H-FL1-
26	704	50.8	749	7	CK982692	4115331.B
27	698	50.4	696	5	BU302606	603739448
28	692	50.0	579	7	CK819035	CK819035
29	691.5	49.9	789	5	BU242187	603781023
30	683	49.3	909	5	BX756548	BX756548
31	682	49.2	548	7	CO880741	CO880741
32	671	48.4	682	2	BB043703	BB043703
33	668	48.2	706	5	BU202465	603949052
34	666	48.1	869	5	BX754527	BX754527
35	647	46.7	919	7	CF407294	CH3H047_E
36	646	46.6	878	5	BX767940	BX767940
37	645	46.6	533	4	BM126453	BM126453
38	638	46.1	1082	5	BX359023	1f06d05.Y
39	625	45.1	506	2	BF415802	UI-R-CAL-
40	618.5	44.7	879	5	BU256052	BU256052
41	614	44.3	861	5	BU246158	603746167
42	611	44.1	1171	5	BU261605	603501763
43	604	43.6	424	1	AA273921	vb99c03.r
44	602	43.5	413	1	AA589724	v178c10.r
45	600	43.3	863	5	BU376295	603808890

ALIGNMENTS

RESULT 1
LOCUS BQ423651
DEFINITION AGENCOURT 7790948 NIH_MGC_72 Homo sapiens CDNA clone IMAGE:6065828
5' mRNA Sequence.
BQ423651
BQ423651.1 GI:2118966
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human).
EST.
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-riemail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LNLN1342 row: 1 column: 21
High quality sequence stop: 710.
Location/Qualifiers
1. 870

FEATURES

source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6065828"

/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NH1 MGc 72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 1.9e-136 Length: 870
Score: 1263.00 Matches: 248
Percent Similarity: 98.42% Conservative: 1
Best Local Similarity: 98.02% Mismatches: 2
Query Match: 91.19% Indels: 2
DB: 5 Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x BQ423651 (1-870)

QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnValGluPheProPheLeuVal 20
Db 112 ATGGAGGCTCATCTACTATTAGTTGACCAATTAAGTAGAATTTCTTTTAGTT 171
QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
Db 172 CTTTGATTCTCGAGGTCACTGCTGTCGCACTTACAGGAGTTTCAGATTCTG 231
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 60
Db 232 CTCTCTGGAAAGCTTTGGACATAGCACAGGTGACATGCTTGCACAGGTGGCAAGAGA 291
QY 61 LeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyValIleGlnHis 80
Db 292 CTTTCTTTTAAATCACTCCAGAGTCTCCACATGATGAGTGGGAAAGCCATGAACT 351
QY 81 LeuAlaLysGlnGlyAsnArgPheHisAspIleLysProProLeuHisHisAlaLys 100
Db 352 TTGGCCAAACAGAAATATGATTTTCATTTGACATCAACCTCCCTTGCATCATGCTAAA 411
QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLys 120
Db 412 AATGTGATTTTCTTTTACTGAGCTTCAACACCTTACTATTAATAATATATGAAAAG 471
QY 121 GluLysGlnGluGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 140
Db 472 GAAAAAGAGAGATTTGAAAGGGCAATCTGCTTCAGCAGCAGCATGCTGCC 531
QY 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
Db 532 ACGATACAGACACAAATGCGATGTCATCTGTGAAAAGAACACATCGGCTATTCTGTT 591
QY 161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyVal 180
Db 592 TGTAAACAGAGACCTTGTACCTCAAAATATATGCACTACTGCTGATCTGGGTGTC 651
QY 181 AlaSerAspPheThrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThr 200
Db 652 GCAAGTAACCTCATATATCCAGAGCTCTGGAATTTTAAACAAACGACACAGTGCAT 711
QY 201 LeuLeuCysProProProArgLeuCysThrAspAsnGlyIleMetIleAlaTyrAspGly 220
Db 712 TTGGTGTCTCTCTCCCAACATATGACATGATATAGCACTTATGATTCAGTGAATGCT 771
QY 221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGlnGlyIleArgTyrGlu 240
Db 772 ATTGAAGAAGCTACTGCTGCTGGCATTTTACATACACATAGAGGATCCGCTATGAA 831
QY 241 -ProLysCysPro-LeuGlyValAspIleSerLys 251
Db 832 ACCAAATGTCTCTCTGGAGTATGACATATCAAAA 866

RESULT 2
AK045669LOCUS
DEFINITION

AK045669 2284 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male corpora quadrigemina cDNA, RIKEN
full-length enriched library, clone:B230219017 product:similar to
PUTATIVE SIMULGYCOPROTEINASE TYPE 2 [Homo sapiens], full insert
sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK045669
AK045669.1 GI:26337528
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus

REFERENCE

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN,
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>

FEATURES

source

1. .2284
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM:DB:B230219017"
/db_xref="taxon:10090"
/clone="B230219017"
/sex="male"
/tissue_type="corpora quadrigemina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
221. .1465
/note="unnamed protein product; putative
similar to PUTATIVE SIMLOGICOPROTEASE TYPE 2 (Homo
sapiens) (SPTR|Q9H4B0, evidence: PASTY, 80%ID, 100%length,
match=1242)"
/codon_start=1
/protein_id="BAC32450.1"
/db_xref="GI:26337529"
/translation="MLMRRTAGALPKPKSKVYGFRRPSVHRPTSCIKLVIGRT
SCDDTGAAYVDENGVLEALHSCVQVHITGCVPPVQAQVRENIQVREYLSAC
RITPSDLASATTKIKPLASLVGSLFSQVNVQKPEPIHMHSAHLTKLNK
VEPPLVLALISGHCCLALVQGVSLPLKSLDIPADMLDKVARELSIKHPECT
MSGKALIEQAKDGNRFTINPMQAKCDSFETGLQHTDYLITHEKEEIEKG
OLSSADIDIAA0VHATFAGHAKRTHRAIFCKOKNLSIPANAYLVSGVASYLYR
KALEIVANRQCTCLCPPLCTCDNGIMANGIERLRAGLVADVEDIRYERKCL
GVDISREVAEAIKVPRLKVAL"
2262. .2267
/note="putative"
2284
/note="putative"

CDS

polya_signal
polya_site
/note="putative"

ORIGIN

Alignment Scores:

Pred. No.: 2,75e-131 Length: 2284
Score: 1224.00 Matches: 233
Percent Similarity: 92.51% Conservativeness: 14
Best Local Similarity: 87.27% Mismatches: 20
Query Match: 88.38% Indels: 0
DB: 3 Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x AK045669 (1-2284)

QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
DB 662 ATGAGAGCTCAGCAGCAGCTATTAAGCTCACCAATAAGTAGAATTTCTTTTAACTT 721
QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
DB 722 CTTTGTATTTCTGGCGGCTCACTGCTGTGGCATTAGCCCAAGGGTTCGATTTCCG 781
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
DB 782 CTCCTTGGAGAGCTTTTGACATAGCACACGACGACATGCTTGACAGGTGGCAAGAGA 841
QY 61 LeuSerLeuIleLysHisProGlyCysSerThrMetSerGlyGlyLysAlaIleGlnHis 80
DB 842 CTTTCTTATTAACAACATCCAGATGTTCTTACATGAGTGCGAAGAGCTATAGAACAG 901
QY 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLys 100
DB 902 TTGGCACAAGACGGAATAATGATTCATTTTACTATCAATCCACCTATGACAGATCTAAG 961
QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
DB 962 AATTGCAATTTTCTTTCACGAGACTTCACATATTAAGTAAATACACACAAAG 1021

QY 121 GluLysGluGluGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 140
DB 1022 GAAAGAGAGAGAGCATTGAGAGAGGGCAATCTGTGCATCAGCGCAGACATGCTCTCT 1081
QY 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
DB 1082 GCGGTACGACATGACACAGCGCTGCACCTTGCAGAAAGACACATCGCGCTATTTGTTT 1141
QY 161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyVal 180
DB 1142 TGCAGAGCAAGAAATTTGCTCTCTCCAGCTAACCCAGATTAAGTTGATTCGAGAGTGT 1201
QY 181 AlaSerAspPheYrIleArgArgAlaLeuGluIleLeuThrAsnAlaIleThrGlnCysThr 200
DB 1202 GCAAGTAACTTGATCATTCGAAAGAGCATTGGAATTCGCAAAATGCAACCCAGCGACAG 1261
QY 201 LeuLeuCysProProProArgLeuLeuCysThrAspAsnGlyIleMetIleAlaIlePheAsn 220
DB 1262 TTGTTGTGTCCACCTCCCAAGACTGTGCATGACATGCAATGCAATTCATGGAATGGA 1321
QY 221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrglu 240
DB 1322 ATGAAAGATTACGTCGCTGCGCTTGGCGCTTTTACATGATGTAAGACATCCGATGAA 1381
QY 241 ProLysCysProLeuGlyValAspIleSerLysGluValGlyGluAlaSerIleLysVal 260
DB 1382 CCAAAATCTCTCTTGAGATGACATATTCAGAGAGTTCGAGAGAGCTCCATAAAGTA 1441
QY 261 ProGlnLeuLysMetGluIle 267
DB 1442 CCGGATTAATAAATGACACTT 1462

RESULT 3

AK011265

LOCUS

DEFINITION

Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:2610001M19 product:similar to PUTATIVE
SIMLOGICOPROTEASE TYPE 2 (Homo sapiens), full insert sequence.

ACCESSION

AK011265.1 GI:12847275

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLES

JOURNALS

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLES

JOURNALS

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLES

JOURNALS

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLES

JOURNALS

MEDLINE

PUBMED

REFERENCE

AUTHORS

1 Carninci, P., and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kikunaka, T., Tashiro, H., Itoh, M.,
Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-394-Format
sequencing pipeline with 384 multichannel sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS
4	THE RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.	Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)		5	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) (bases 1 to 1622)			
		Aadachi,T., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukanishi,Y., Furuno,M., Hanagaki,T., Harai,A., Hayatsu,N., Hiramoto,K., Hirotsuka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kikukawa,T., Kato,H., Kawai,Y., Kojiima,Y., Komoto,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishii,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Saeki,D., Shibata,K., Shioda,Y., Shinigawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toyai,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.			
		Direct Submission			
		Submitted (10-UTL-2000) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)			
		Please visit our web site (http://genome.gsc.riken.jp/) for further details.			
		cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGACGATCCAGACCTCTTTTCTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTTCGATTAATTAATATCATCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.			
FEATURES	SOURCE	location/Qualifiers			
		1..1622			
		/organism="Mus musculus"			
		/mol_type="mRNA"			
		/strain="C57BL/6J"			
		/db_xref="FANTOM DB:261000IM19"			
		/db_xref="taxon:10090"			
		/clone="261000IM19"			
		/tissue="whole body"			
		/clone_id="RIKEN full-length enriched mouse cDNA library"			
		/dev_stage="10 days embryo"			
		207..1451			
		/note="unnamed protein product; putative similar to PUTATIVE SIGNAL PROTEIN TYPE 2 (Homo sapiens) (SPTR Q9H4B0, evidence: FASTV, 80%ID, 100%length, match=1242)"			
CDS		/codon_start=1			
		/protein_id="BAB7506.1"			
		/db_xref="GI:12847276"			
		/translation="MTMLRRKAIIPRPSSKYGLPRLRSYVPRITSGHVLGIETSCDPTGAADVDETNVLGEALSTGYQLVTGVPPAQDLHENTQRVEETLSACRISDSDIAATITKPGIALSIGVLFSLQIVNQPKPPIPIHMEASHALTITLRKVSEFPFLILSGCHLAVOGVDPLFGSLDIADMDKDVARLSLEIKKECGTMSGGKALEOLAKGNRHFTINPMONAKNDPSFGTIOHTDGLITHKEEGIEKKRQIISSADDIAAIVOHATRAHLARTTHAILFCCKNLLSPANNLVVSGYSNLTYIRRKALIVANATCTCLCPRLCTDNGIMTIANGIERLAGVGVLHVEDIDRYEKPCPLGVDSIRSRVAEALIVPRLKVAL"			

polysignal	1605..1610	/note="putative"
polysite	1622	
polysite	1622	/note="putative"
ORIGIN		
Alignment Scores:		
Pred. No.:	1,44e-130	Length: 1622
Score:	1216.00	Matches: 232
Percent Similarity:	92.13%	Conservative: 14
Best Local Similarity:	86.89%	Mismatches: 21
Query Match:	87.80%	Indels: 0
DB:	3	Gaps: 0
US-10-649-273-2_COPY_148_414 (1-267) x AK011265 (1-1622)		
QY	1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGlnPheProPheLeuVal	20
Db	648 ATGAGAGGTCAAGCAGCATGACTATTAGGCTCACCAATAAGTAGAATTCTTTTAGTT	707
QY	21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu	40
Db	708 CTTTGGATTCTGGGGGTCACTGGCTGTGGGATTAAGTCCAAAGGTGTTCCGATTTCCCTG	767
QY	41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgAsp	60
Db	768 CTCCTTGGGAAGTCTTTGAGCATACACAGCGACATGCTTGAACAAGGTGGCAAGA	827
QY	61 LeuSerLeuIleLysHisProGlnCysSerThrMetSerGlyGlyLysAlaIleGlnHis	80
Db	828 CTTTCTTTAATCAAAATCCAGAAAGTCTACAAATGAGTGGTGAAGAAAGCTATAGAACG	887
QY	81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisAlaLys	100
Db	888 TTGGCAAAAGACGGAATAGATTCCATTTATCTATCAATCCACTATGCAAGATGCTAAG	947
QY	101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys	120
Db	948 AATTGCGATTTTCTTTCAACGGGACTTCAACATATTACTGATAGGCTAATATACACACAG	1000
QY	121 GluLysGlnGlnGlyIleGlnLysGlnIleLeuSerSerAlaAlaAspIleAla	140
Db	1008 GAAAAGAGAGAGCATTTGAGAGAGCAAACTCTGTATCAGCTGCACACATGCTGCT	106
QY	141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe	160
Db	1068 GCGGTACAGCATGCAACAGCGTGCACCTTGGAAAAGAACATGCGCCTATTCTGTTT	112
QY	161 CysLysGlnArgAspLeuLeuProGlnAsnAlaValLeuValAlaSerGlyGlyVal	180
Db	1128 TGCAAGACGAAAATTTTCTCTCTCCACCTAACGGATTAAGTGTATCTGGAGGTGT	118
QY	181 AlaSerAsnPheTyrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThr	200
Db	1188 GCAAGTAACTTGATCATCCGAAGACATTTGGAATTTGTGCGAAATGCAACGCAGTGCAG	124
QY	201 LeuLeuCysProProProArgLeuCysThrAspAsnGlyIleMetIleAlaIleAsnGly	220
Db	1248 TTGTTGTGTCCACCTCCAAAGACTGTGCACTGACATGGCATCATGATTCATGGAATGGA	130
QY	221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGln	240
Db	1308 ATTTGAAGATTAACTGCTGGCTGGCGCTGTTTACATGATGATGAGACACATCCCATATGAA	136
QY	241 ProLysCysProLeuGlyValAspIleSerLysGlnValGlyGluAlaSerIleLysVal	260
Db	1368 CCAAAATATCTCTTGGAGTACATATTCACAGAGAAATTTGCAAGAGCTGCATATAAGTA	142
QY	261 ProGlnLeuLysMetGluIle 267	
Db	1428 CCGCGATTAAATAATGCACTT 1448	

B0636028
 LOCUS B0636028 640 bp mRNA linear EST 15-JUN-2002
 DEFINITION hd03d11.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
 Homo sapiens cDNA clone hd03d11 5', mRNA sequence.
 ACCESSION B0636028
 VERSION B0636028.1 GI:21760487
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 640)
 Wistow,G., Bernstein,S.T., Wyatt,M.K., Ray,S., Behal,A.,
 Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
 Expressed sequence tag analysis of human retina for the NEIBank
 Project: Retbindin, an abundant, novel retinal cDNA and alternative
 splicing of other retina-preferred gene transcripts
 Mol. Vis. 8 (4), 196-204 (2002)
 JOURNAL 22103461
 MEDLINE 12107411
 PUBMED
 COMMENT Contact: Wistow G
 Section on Molecular Structure and Function
 National Eye Institute
 6/331, NIH, Bethesda, MD 20892-2740, USA
 Tel: 301 402 3452
 Fax: 301 496 0078
 Email: graeme@helix.nih.gov
 Plate: 03 row: d column: 11
 Seq primer: M13RP1 reverse primer (ABI).
 Location/Qualifiers
 1..640
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="hd03d11"
 /tissue_type="Retina"
 /dev_stage="Adult"
 /lab_host="EMDH10B"
 /clone_lib="Human Retina cDNA (Un-normalized,
 unamplified): hd/he"
 /note="Organ: Eye; Vector: pSPORT1; Neutral retina tissue
 was dissected from two 80 year old donors with no observed
 eye disease. 100ug of total RNA was used for library
 construction. A directionally cloned cDNA library in the
 pSPORT1 vector (Life technologies) was constructed at
 Bioserve Biotechnology (Laurel MD) essentially following
 the protocols of the Superscript Plasmid System full
 details of which are contained in the manufacturer's
 instruction manual (http://www.lifetech.com/). First
 strand synthesis was carried out using a Not I
 primer-adaptor
 5'-PACTAGTTCTAGATCGAGCGCGCCGCC(7)15-3'. EST analysis
 was performed on the unamplified library at the NIH
 Intramural Sequencing Center (NISC)."
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,94e-117 Length: 640
 Score: 1097.00 Matches: 209
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 79.21% Indels: 0
 DB: 5 Gaps: 0
 US-10-649-273-2_COPY_148_414 (1-267) x B0636028 (1-640)
 Oy 59 ArgarglseuSerleuilelyshisproglucysseerthetsergllygylvalaile 78
 Db 3 AGAAGACTTTCTTAATTAACATCCAGAGGCTCCACATGATGGAGGAAAGCCATA 62
 Oy 79 GtuhileuAlalygsginglyasnargPhehisPheapillelyspptroleuhis 98
 Db 63 GAACATTTGGCCAAACAAAGAAATAGATTTCATTTGACATCAACCTCCCTGCATCAT 122

Oy 99 AlalyshancysaapPheSerPheThGlyleuglnhisvalthrasplylelle 118
 Db 123 GCTAAATATGTGATTTTCTTTACTGACCTTCAACGCTTACGATTAATATATG 182
 Oy 119 lyelsgsluylusgluglylileglulysgllylnileuserSerAlaAlasplle 138
 Db 183 AAAAAAGAAAAAGGAAAGGATTAAGAAAGGGCAATCTGTCTTCAGCAGCAGACAT 242
 Oy 139 AlAlathrValglnhsthMetAlaCyshisluvallylsargThhisargAlaile 158
 Db 243 GCTCCACAGTACGACACACATGCAATGCAATCTCTTGAAAGAAACATCGGCTATT 302
 Oy 159 leuPheCyslysglnargaspPleuleProglinsnaenalavalleuvalasergly 178
 Db 303 CTGTTTGTAAAGAGAGAGACTGTACTCAAAATATGAGTACTGGTGCATCTGCT 362
 Oy 179 GlyValAlaseranPheTyllleargArgAlaleuglulleuthThasnaAlathrgln 198
 Db 363 GGTCTCCAGTAACTTCTATATATCCGACAGCTCTGAAATATTTAAACAGCACACAG 422
 Oy 199 CyethrleuPheCysproProargleuCyethrAspanglylMetllelathp 218
 Db 423 TGCACTTGTGTGTCTCTCCACAGACTATGACTGATTAATGGATTAATGATTCATGG 482
 Oy 219 AasnGlyllegluAqgleuargAlaglyleuglylleuHisaspillegluglyllearg 238
 Db 483 AATGATTTGAAGAAACATTCGTCGCTGCGCATTTACATGACATGAAGCATCCGC 542
 Oy 239 TyrgluProlyscysproleuglyvalasplleserlysglyvalaglyluAlaserlle 258
 Db 543 TATGACCAAAATATGCTCTTGAGTACATATCAAAAGAAAGTTGGAAGCTTCCATA 602
 Oy 259 lvsValProglinleuylsMetglulle 267
 Db 603 AAGTACCACAAATTAATAATGAGATA 629
 RESULT 5
 EX391919/c 852 bp mRNA linear EST 28-APR-2004
 LOCUS EX391919 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DB001YE02 3-PRIME, mRNA sequence.
 ACCESSION EX391919
 VERSION EX391919.2 GI:46846154
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 852)
 Li,W.B., Gruber,C., Jesse,J., and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 13, 2003 this sequence version replaced gi:30611736.
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 1240.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdnas=CS0BA1035ZB07_CS03317_1fc=1240.r
 FEATURES
 source
 location/Qualifiers
 1..852
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DK001YB02"

```

/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_idb="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/Note="1st strand cDNA was primed with a NotI-oligo(5')
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Alignment Scores:

```

Pred. No.: 1 65e-114 Length: 852
Score: 1075.00 Matches: 208
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 77.62% Indels: 0
DB: 5 Gaps: 0

```

US-10-649-273-2_COPY_148_414 (1-267) x BK391919 (1-852)

```

QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnValGluPheProPheLeuVal 20
Db 626 ATGGAGGCTCATCTACTACTATTAGCTTGAACATAAGATTTCTTTTATGTT 567
QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
Db 566 CTTTGTATTCTTGAGGCTCATCTCTGTTGGCATTTAGTCAAGAGTTTCAGATTCTG 507
QY 41 LeuLeuGlyLeuSerLeuAspIleAlaProGlyAspMetLeuAspValAlaArg 60
Db 506 CTTCTTGGAAGGCTTTGACATAGCACAGGTCACATGCTTGAACAAGTGGCAAGAGA 447
QY 61 LeuSerLeuIleHisHisProGlyCysSerThrMetSerGlyGlyValIleGlnHis 80
Db 446 CTTTCTTTAATAAATCCAGAGTGTCTCAACCATGATGTTGGAGAAAGCATGAAAT 387
QY 81 LeuAlaValGlnGlyValAsnArgPheHisPheAspIleLeuAspProLeuHisHisAlaVal 100
Db 386 TTGGCCAAACCAAGAAATAGATTCTTTCATTCATCAACCTCCCTTGACATCATGATAA 327
QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspValIleIleMetLeuLys 120
Db 326 AATTGATTTTCTTTTACTGACATTCACACCTTCTGTAATAAATAAATGAAAAAG 267
QY 121 GluValGlnGluGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAlaAla 140
Db 266 GAAAAAGAGAGAGATTTGAAAGGGCAATCTGTCTTCAGCAGCAGACATTTGCTGCC 207
QY 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
Db 206 AAGGTCACAGACACATGATGATCTTGTGAAAGAACACATCGGGCATTTCTGTTT 147
QY 161 CysGlyGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyVal 180
Db 146 TGTAAACAGAGACATCTTACCTCAAAATAATGACATGCTGATGCTGCTGCTC 87
QY 181 AlaSerAspPheThrIleArgArgAlaLeuGlnIleLeuPheAsnAlaThrGlnCysThr 200
Db 86 GCAAGTAACCTTATATCCGACAGGCTCTGAAATTTTAAACAAACGACACATGACACT 27
QY 201 LeuLeuCysProProProArgLeu 208
Db 26 TTGTTGTGTCCTCTCCACAGACTA 3

```

```

RESULT 6
BC030671 1609 bp mRNA linear HTC 19-NOV-2003
LOCUS BC030671 Mus musculus O-sialoglycoprotein endopeptidase-like 1, mRNA (CDNA
DEFINITION clone IMAGE:1226118), containing frame-shift errors.
ACCESSION BC030671
VERSION BC030671.1 GI:21040459
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

```

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionath; Muridae; Mus. 1 (bases 1 to 1609)

Klausner, R.D., Collins, F.S., Wagner, L., Sherman, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buell, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stappleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Tothiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mulvaney, S.J., Bosak, S.A., McKean, P.J., McKernan, R.D., Malek, U.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A.C., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalys, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 1609)

Strausberg, R.

Direct Submission

Submitted (20-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgaps-remail.nih.gov

Tissue Procurement: Marcello Bento Soares, Ph.D.

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: amadansystemsbio.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Series: IRAX Plate: 66 Row: e Column: 10

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21312463

This clone has the following problem: frame shifted.

FEATURES

source

```

1..1609
Location/Qualifiers
  /organism="Mus musculus"
  /mol_type="mRNA"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="IMAGE:1226118"
  /tissue_type="Thymus gland, mouse"
  /clone_idb="Soares_thymus_2ndwt"
  /lab_host="DH10B"
  /note="Vector: pT73-Pac"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 1 21e-113 Length: 1609
Score: 1071.00 Matches: 211
Percent Similarity: 83.52% Conservative: 12
Best Local Similarity: 79.03% Mismatches: 20
Query Match: 77.33% Indels: 24
DB: 3 Gaps: 1

```

US-10-649-273-2_COPY_148_414 (1-267) x BC030671 (1-1609)

QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
Db 665 ATGGAGGCTACCGACGACGATATTAGGCTCACCAATAAGTAATTTCTTTTACTT 724
QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
Db 725 CTTTGTGATTTCTGGCGGCTGACTGCTGTGGCATTTAGTCCAAAGGTGTTTCGATTTCTGT 784
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 60
Db 785 CTCTTGGGAGGCTTTGGACATAGCACAGGCGACATGCTTGCAAGGTGCAAGAGA 844
QY 61 LeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyValAlaIleGluHis 80
Db 845 CTTTCTTATTAACAACATCCAGATGTTTCAATGATGCTGAAAAAGCTATAGAAG 904
QY 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLys 100
Db 905 TTGGCCAAAGACGGAATAGATTCATTTTACTATCAATCACTAGCAAGATGCTAAG 964
QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
Db 965 AATTGCAATTTTCTTTCACGGGACTTCAACATTTACTGATAGCTAATAACACACAG 1024
QY 121 GluLysGluGluGlyIleGlyLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaAla 140
Db 1025 GAAAAAAGAGAGGACTTGGAGAGGGGCAAACTGTCATAGCTGACAGCATTTGCTCT 1084
QY 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
Db 1085 GCGGTCACGATGACACAGCGTGCACCTTGCAAAAGAACATCGCGCTATTTCT 1144
QY 161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyVal 180
Db 1145 TGCAGACGAGAAAATTTGCTCTCTCCAGCTAACCCAGATTTAGTTGATCGAGGTCT 1204
QY 181 AlaSerAsnPheThrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCysThr 200
Db 1205 GCAAGTACTTGTATCATCCGAAAGCATGTGCAAAATGTGCAAAATGCAACGACAGT 1264
QY 201 LeuLeuCysProProArgLeuCysThrAspAsnGlyIleMetIleAlaIleArgAsnGly 220
Db 1265 TTGTTGTCCTCCTTGGAGATATCCAGAGAGTTGCAAGAGCTGCATTAATAAGTA 1372
QY 221 IleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgGly 240
Db 1315 ----- 1315
QY 241 ProLysCysProLeuGlyValAspIleSerLysGlyValGlyGluAlaSerIleLysVal 260
Db 1316 ---TGAGTCTCTTGGAGATATCCAGAGAGTTGCAAGAGCTGCATTAATAAGTA 1372
QY 261 ProGlnLysMetGluIle 267
Db 1373 CCGGATTAATAAATGCACTT 1393
RESULT 7
BE740611 701 bp mRNA linear EST 15-SEP-2000
LOCUS 601595739P1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3949640 5',
DEFINITION mRNA sequence.
ACCESSION BE740611
VERSION BE740611.1 GI:10154603
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 701)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strusberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DBP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM814 row: n column: 09
High quality sequence stop: 701.
Location/Qualifiers
1..701
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3949640"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_9"
/note="Organ: ovary; Vector: pOTB7, Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN
Alignment Scores:
Pred. No.: 4.6e-112 Length: 701
Score: 1053.00 Matches: 206
Percent Similarity: 99.05% Conservative: 3
Best Local Similarity: 97.63% Mismatches: 0
Query Match: 76.03% Indels: 2
DB: 2 Gaps: 0
US-10-649-273-2_COPY_148_414 (1-267) x BE740611 (1-701)
QY 1 MetGluAlaHisAlaLeuThrIleArgLeuThrAsnLysValGluPheProPheLeuVal 20
Db 71 ATGGAGGCTACCGACGACGATATTAGGCTCACCAATAAGTAATTTCTTTTACTT 130
QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
Db 131 CTTTGTGATTTCTGGCGGCTGACTGCTGTGGCATTTAGTCCAAAGGTGTTTCGATTTCTGT 190
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 60
Db 191 CTTTCTTGAAGACTCTTTGACATACACAGGTGACATGCTTGACAGGTGCAAGAGA 250
QY 61 LeuSerLeuIleLysHisProGluCysSerThrMetSerGlyGlyValAlaIleGluHis 80
Db 251 CTTTCTTAAATTAACATCCAGAGGTCTCCACATGAGAGTGGGAAAGCCCTAGAACAT 310
QY 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLys 100
Db 311 TTGGCCAAAGACGGAATAGATTTGATTTGACATCAAACTCTTCGATCATGCTAAG 370
QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
Db 371 AATTGCAATTTTCTTTCACGGGACTTCAACATTTACTGATTAATAATAAGTAAGTA 430
QY 120 YGluLysGluGluGlyIleGlyLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 140
Db 431 GGAAGACAGAGAGATTTGAGAGGGGCAAACTGCTTCACAGACAGCATTTGCTG 490
QY 140 IeThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuP 160
Db 491 CCACAGTACAGACACATAGCATGCTCTGTGTAAGAAACACATCGGCTATTTCTCT 550
QY 160 HeCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGly 180

Db 551 TTGTGACGACAGAGACTTGTACCTCAAAATAATGACATGCTGTCATCTGCTGTC 610
Qy 180 a1a1aseaenpethy71leargalaleu9guilleuThrasnAlathrGlnCysT 200
Db 611 TCCAGTACTTCTATATCCGACAGCTCTGGAATTTTACAAAGCAACAGATGCA 670
Qy 200 hrleuLeuCySPProProProargLeuCyS 209
Db 671 CTTGTGTGTCCTCTCCAGACTATGC 699
RESULT 8
BQ961028 922 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT 8863711 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6423902
DEFINITION 5', mRNA sequence.
ACCESSION BQ961028
VERSION BQ961028.1 GI:22376506
KEYWORDS EST.
SOURCE Homo sapiens (human).
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 922)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Plate: LLCM2603 row: d column: 15
High quality sequence stop: 584.
location/Qualifiers
1..922
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6423902"
/issue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_18"
/note="Organ: lung; Vector: pOT7; Site: 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA Synthesis Kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Alignment Scores:
Pied. No.: 3.65e-99 Length: 922
Score: 944.00 Matches: 207
Percent Similarity: 89.45% Conservative: 5
Best Local Similarity: 87.34% Mismatches: 16
Query Match: 68.16% Indels: 10
Gaps: 3
DB: 5
US-10-649-273-2_COPY_148_414 (1-267) x BQ961028 (1-922)
Qy 1 MetGlAlaHisAlaLeuThrIleArgLeuThrasnysValGluPheProPheLeuVal 20
Db 207 ATGAGAGCTATGACCTACTTACTATTAGGTGACCAATAGAGATTCTTTTATGATT 266
Qy 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValIserAspPheLeu 40
Db 267 CTTTGAATTTCTGAGAGTCACTGTCCTTGCAATTAGTTCAGAGATTTCGATTCTG 326

Qy 41 LeuLeuGlyIysSerLeuAspIleAlaProGlyAspPheLeuAspIysValAlaArgArg 60
Db 327 CTTCTTGAAAGGTCTTTGACATAGCACAGGTGACATGCTTGAACAAGTGGCAAGAGA 386
Qy 61 LeuSerLeuIleIysHisProGlyCys-SerThreSerGlyGlyValaIleGluHis 80
Db 387 CTTTCTTTAATMAAACATCCAGAGTGCNTCCACCATGATGTGTGGAAAGCCATAGAAC 446
Qy 80 sleuAlaIysGlnGlyAsnArgPheHisPheAspIleIysProProLeuHisAlaIly 100
Db 447 TTGGCCAAACAGCAAAATGATTTTCATTTTGACATCAAACTCCCTTGCAATCATGCTAA 506
Qy 100 sasnCyAspPheSerPheThrgIyLeuGlnHisValThrAspIysIleIleIysIly 120
Db 507 AAATGTGATTTTCTTTTACATGACCTTCAACAGCTTACTATTAATAATATGAAAAA 566
Qy 120 sgluIysGluGluGlyIleGlyIysGlyGlnIleLeuSerSerAlaIaAspIleAlaI 140
Db 567 GGAATAAGAGGAAAGATTTGAGAAAGGCAAAATCCTGTCTTCAGCAGACGATTCGTC 626
Qy 140 aThrValGlnHisThrMetAlaCysHisleuValIysArgThrHisArgAlaIleLeuPh 160
Db 627 CACAGTACAGCACCAATAGCATGTCATCTGTGAAAAAGAAA-CATCGGGCTATTCTGTT 685
Qy 160 eCyAlysGlnArgAspLeuLeuProGlnAsnAsnAlaValIleuValAlaser-GlyIly 180
Db 686 TTGTAACGACAGAGACTTGTGACTCAAAATATATGCACTATGGGTGATCTGGGGGG 745
Qy 180 a1a1aseaenpethy71leargalaleu9guilleuThr-AsnAlathrGlnCys 199
Db 746 TCCAGATTAATCTTATATCCCGCAACTGGAATAATTACAAACCAACGAGGGGC 805
Qy 200 Thr-LeuLeuCySPProProProargLeuCyThr-AspAsnGlyIleMetIleAlaTrp- 218
Db 806 ACCTTGTGTGGTCCCTCCCAACTATGACGAGATTAATGGG---CATTAATGATGTC 862
Qy 219 -----AsnGlyIleGluArg---LeuArgAlaGlyLeuGly 229
Db 863 TGGGAAAGGAAATGAAAAAATACATCTGCTCGGCTGGGG 903
RESULT 9
CF362328 658 bp mRNA linear EST 25-AUG-2003
LOCUS CF362328/c
DEFINITION 829596 MARC 3Pig Sus scrofa cDNA 3', mRNA sequence.
ACCESSION CF362328
VERSION CF362328.1 GI:34161882
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
REFERENCE 1 (bases 1 to 658)
AUTHORS Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.T., Fox,J.,
Wise,T.A., Noneman,D.J., Wray,J.E. and Keeler,J.W.
TITLE A second set of porcine ESTs from a pooled-tissue normalized
library
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4350
Email: smitht@mail.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: SRG8015 row: H column: 12
Seq primer: TAGAAGGACAGCTCAGG.
FEATURES
source
1..658
/organism="Sus scrofa"
/mol_type="mRNA"

/db_xref="taxon:9823"
/issue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 3Pig"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including brain, liver, muscle, placenta/endometrium,
ovary, testes, and bone marrow."

ORIGIN

Alignment Scores:
Pred. No.: 3,71e-97 Length: 658
Score: 925.00 Matches: 174
Percent Similarity: 93.97% Conservative: 13
Best Local Similarity: 87.44% Mismatches: 12
Query Match: 66.79% Indels: 0
DB: 7 Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x CF362328 (1-658)

QY 20 ValLeuLeuLeuLeuSerGlyGlyHisCysLeuLeuValGlnGlyValSerAspPhe 39
DB 657 GTTCTTTGATATCTGGCGGTCATTTCTTTGGCATTTAGAGAGATTTCAGATT 598
QY 40 LeuLeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArg 59
DB 597 CTGCTTCTTGACAGCTTTGGACATAGCACAGCTGACATGCTTGACAGAGTGGACAA 538
QY 60 ArgLeuSerLeuLeuLysHisProGlyCysSerThrMetSerGlyGlyLysAlaIleGlu 79
DB 537 AGACTTTCTTTAATAAATCCAGAGTGTCTCCACATGATGTGGGAGGCAATAGAA 478
QY 80 HisLeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisAla 99
DB 477 CATTTGTCACAAACAGGAAATTAAGTTCATTTGATTTCAAACTCCCATGCAACGTGAT 418
QY 100 LysAsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLys 119
DB 417 AAAAATGTATTTTCTTTTCTGAGCTTCACAGTTATTGATAGACATTAATGCG 358
QY 120 LysGlnLysGlnGlyGlyIleGlnLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 139
DB 357 AAGGAAAGAGAGAGATTTGAGAGGGGCAATCTGCTTCAGCTGACAGATTCT 298
QY 140 AlaThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeu 159
DB 297 GCTGCACTACAGACACAGTACCTGCGCATATTCACAAAGAACTCATCGTATTCTG 238
QY 160 PheCysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGly 179
DB 237 TTTTGGCAACAGAGAGCTTATTTATGTCAAAGTAATGCAATGCTGTATCTGAGAGT 178
QY 180 ValAlaSerAsnPheTyrlleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCys 199
DB 177 GTTGCAAGTAACTTATATATACAAAGCTTTAAGAGTTTGACAAATGCAACAGATGC 118
QY 200 ThrLeuLeuCysProProArgLeuCysThrAspAsnGlyIleMetIleAlaIleTyr 218
DB 117 ACTCTGTGTGTCTCTCCAGACTATGCACTGATATGCAATTATATGTCATG 61

RESULT 10
CF257246 822 bp mRNA linear EST 07-AUG-2003
LOCUS Phn008 g02 PHA-activated splenocytes Gallus gallus cDNA, mRNA
DEFINITION sequence.
ACCESSION CF257246
VERSION CF257246.1 GI:33490501
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 822)
AUTHORS Wiltzell,H., Bed'Hom,B., Morin,V., Young,J.R., Whitaker,C.J.,
Chausse,A.M. and Zoorob,R.
TITLE A collection of chicken ESTs from activated immune cells
JOURNAL Unpublished (2003)
COMMENT Contact: Zoorob R
UPR 1983
CNRS

7 rue Guy Moquet, BP 8, 94801 Villejuif cedex, France
Tel: 33 1 49 58 35 00
Fax: 33 1 49 58 33 81
Email: zoorob@vjf.cnrs.fr.
Location/Qualifiers

FEATURES

source 1..822
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/cell_type="Splenocytes"
/clone_lib="PHA-activated splenocytes"
/note="Vector: pTriplEx2"

ORIGIN

Alignment Scores:
Pred. No.: 5,75e-96 Length: 822
Score: 916.00 Matches: 175
Percent Similarity: 81.37% Conservative: 39
Best Local Similarity: 66.54% Mismatches: 49
Query Match: 66.14% Indels: 1
DB: 7 Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x CF257246 (1-822)

QY 1 MetGlnAlaHisAlaLeuThrIleArgLeuThrAsnLysValGlnPheProPheLeuVal 20
DB 35 ATGAGGCTCAGCGCATTAACATCACTGACAGAGCAAGTAATTCCTTCTTAGT 94
QY 21 LeuLeuIleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
DB 95 CTTTACTCTCGGAGAGTCACTGCATCTTGGCAGTAGACACAGAGATTTCATTCCT 154
QY 41 LeuLeuGlyLysSerLeuAspIleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
DB 155 CTGCTTGACAGCTCATGATATACAGCACAGGTGACATGTTGATAGGTAGCAAGAAAG 214
QY 61 LeuSerLeuIleLysHisProGlyCysSerThrMetSerGlyGlyLysAlaIleGlnHis 80
DB 215 CTCTTTTATGAAGACACCGGAGTGCACAGCATGCGGGGGGAGGCAATAGAGAC 274
QY 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisAlaLys 100
DB 275 CTGGCTCAACCGGAGACTGGGCACAGTACACTTTCAGACTTCCCATGCAACAGTATCT 334
QY 101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 120
DB 335 AACTGTATTTTCTTCTTCTCCGACTTCAGAGCTTTCAGCAAAAGCCATTCTTCAGAA 394
QY 121 GlnLysGlnGlyGlyIleGlnLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 140
DB 395 GAAAAAGAAAGAGTATTCAGAAAGGGAATCTGCTCGCTTAAAGACATCCCTGCT 454
QY 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
DB 455 GCTGCAACAGAGTATGCTGCTCATATTTATCCAGCGACACACCGAGCCATGCTCTTC 514
QY 161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyVal 180
DB 515 TGCATGAAAAACAGATATTTATACCAAAATCTGACATCTGCTGTATCAGAGAGATT 574
QY 181 AlaSerAsnPheTyrlleArgArgAlaLeuGlnIleCysThrAsnAlaThrGlnCysThr 200
DB 575 GCAACTATTCAGTATATCAGAAAGAGCTGCACTGCGCAATGCAACGTTTGTCT 634
QY 201 LeuLeuCysProProArgLeuCysThrAspAsnGlyIleMetIleAlaIleTyrAsnGly 220

Db	635	TTTCGTCTCCTCCCTCAAGCGTGGACCCGATTAATGCTGTATTGCATTGCAGGAATGCC	694
Oy	221	leGUatgLeuaRgaLaGlylVeUGLytlleUehisaaplllegUglytllaargtyrGu	240
Db	695	ATTGAAAGGTTCGCGCAGAGATGNGATTTTTATTCACATACTACGATGCGATCCGCTACGAA	754
Oy	241	ProlysCysProlenUglValAspIleSerlySGluValIGlyGUAlASerlleYval	260
Db	755	CCAAAAGCTCCCTTGGAATGATATTTTCCAAGAAGTTGAA-GAAGATTCATCAAAGTG	813
Oy	261	ProGUlenU 263	
Db	814	 CCAAGACTA 822	
RESULT 11			
LOCUS	CK941819/c		
DEFINITION	4065407 BARC 10BOV Bos taurus cDNA clone 10BOV12_F24 3', mRNA	637 bp	mRNA linear EST 15-MAR-2004
VERSION	CK941819		
KEYWORDS	CK941819.1 GI:45456199		
SOURCE	EST.		
ORGANISM	Bos taurus (cow)		
	Bos taurus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
	Bovinae; Bos.		
	1 (bases 1 to 637)		
	Somstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harnay,		
	G.P., Boasek, S., Rubenstein, M. and Gasparre, L.C.		
	Production of EST from cDNA libraries derived from immunologically		
	activated bovine gut		
	Unpublished (2004)		
JOURNAL	Contact: Tad S. Somstegard		
COMMENT	Bovine Functional Genomics Laboratory		
	Animal and Natural Resources Institute		
	Bldg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA		
	Tel.: 3015048416		
	Fax: 3015048414		
	Email: tads@anri.barc.usda.gov		
	Single pass sequencing. Bases called and trimmed with phred		
	0.000925 using options -trim.alt "-" -trim.fasta. Vector identified		
	by cross_match using options -mismatch 12 -minscore 18		
	Plate: 12 row: F column: 24		
	Seq primer: AGCGGATACAATTCACACAGG		
	High quality sequence stop: 637.		
FEATURES			
source	location/Qualifiers		
	1..637		
	/organism="Bos taurus"		
	/mol_type="mRNA"		
	/strain="Holstein"		
	/db_xref="taxon:9913"		
	/clone="10BOV12_F24"		
	/sex="Male"		
	/tissue_type="Pooled"		
	/dev_stage="Multiple"		
	/lab_host="DH10B TI phage resistant"		
	/clone_lib="BARC 10BOV"		
	/note="Organ: Small Intestine; Vector: pagen-1; Site_1:		
	ECOV; Site_2: NOTI; Equimolar amounts of mRNA extracted		
	from proximal jejunums of 18 and 21 wk old steers, and		
	distal ileums of 14 day old calves; proximal jejunum		
	exposed to C. oncophora for 3 and 6 weeks, and distal		
	ileum exposed to C. parvum for 7 days"		
ALIGNMENT SCORES:			
Pred. NO.:	5.09e-87	Length:	637
Score:	838.00	Matches:	153
Percent Similarity:	92.47%	Conservative:	19
Best Local Similarity:	82.26%	Mismatches:	14
Query Match:	60.51%	Indels:	0

DB: 7 Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x CK941819 (1-637)

QY 82 A1a1ysg1ng1ya5na7qphEhi5Pheasp11elysPProleuHi5hi5a1a1ysa5n 101
 Db 632 GCCAAACGAGGAAADTAGTTGATTCATTGTTGATTTCCAGCTCCCAACGCGTAAAAAT 573

QY 102 CysaPpheSerPheThrGlyLeuGlnhi5ValThrAsp1yle11emetyly5Glu 121
 Db 572 TGTGATTTTCTTTTCTGGACTTCAACACGTTTATGATGAAGATATGCAAAAGAA 513

QY 122 LysGluGluGluY11egiulysG1Gln11eleuSerSer1aa1aasp11ealalaThr 141
 Db 512 AAAGGAGAGGATATGACGAGGGACAGGTCTGTCTTACGCTGCAGACATTGCTGGC 453

QY 142 ValGlnhi5ThrMetAlaCyShi5LeuValLysArGthrhi5Argala1leleuPheCys 161
 Db 452 GTGCAGCAGCACCGTGGCTGCCTGCACATTGCAGAAAAGAACACATGCTGCTTCTGCG 393

QY 162 LysGlnArGAsPLeuLeuProGlnAspAsnAlaValLeuValAlaSerGlyValAla 181
 Db 392 AAGCAGAGAGGCTTCTTAAGTCAAGATGACACGCACTGCTGTATGTCGAGGCGTCGCA 333

QY 182 SerAspPheTyr11eArGArGAla1eGlu11eleuThrAsnAlaThrng1n5sThrLeu 201
 Db 332 AGTAACTTATATATCCGAAAGCCCTGGAAATGTGACCAATGCAACACGTCGACTTGG 273

QY 202 LeuCySPProPProArGLeuCy5ThrAspAsnGly11eMet1lealatrPAsnGly1le 221
 Db 272 CTGTGCCCCGCCCCCAAGCTTTGACAGTACAGACGCGCTATGATGCAATGGAATGGTGT 213

QY 222 GluArGLeuArGAlaGlyLeuGly11eleuhi5asp11eGluGly11eArGTYGluPro 241
 Db 212 GAAAGACTACACTGCTGCTGGCTGGCATTTTACACACACAGAGGCAATCCGCTAGGAACA 153

QY 242 LysCySPProLeuGlyValAsp11eSer1y5GluValGlyGluAlaSer11ely5ValPro 261
 Db 152 AAAGTCTCTCTTGAGTGAATATCAAAAGATGTGAGAGCTGTATTAAGTGGCA 93

QY 262 GlnLeuYsMetGluile 267
 Db 92 AGATTAAAAATGAAGATT 75

RESULT 12
 CB272391
 LOCUS CB272391
 DEFINITION mRNA sequence.
 ACCESSION CB272391
 VERSION CB272391
 KEYWORDS Mus musculus (house mouse)
 SOURCE EST.
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus. 1 (bases 1 to 597)
 McCarrey,J., Eddy,M., Marra,M., Hillier,L., Clifton,S., Pape,D., Martin,J., Wylie,T., Dante,M., Bowers,Y., Theisinger,B., Gibbons,M., Ritter,E., Tagarehivilli,R., Ronko,I., Maguire,L., Kennedy,S., Bennett,J., Waterston,R. and Wilson,R.
 NIH5 Mouse
 Unpublished (2002)
 Contact: McCarrey/Bddy NIH5 Mouse
 NIH5 Mouse
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library constructed and donated by J. McCarrey, Ph.D. (Southwest Foundation for Biomedical Research, Dept. of Genetics) - exsiston

done by E.M. Eddy, Ph.D. (National Institutes of Health, National
Institute of Environmental Health Sciences).

MGI:2069710

Seq primer: Primer name ambiguous

High quality sequence stop: 419.

FEATURES

source

Location/Qualifiers

1..597

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CD-1"

/db_xref="taxon:10090"

/clone="IMAGE:6445750"

/sex="male"

/tissue_type="Spermatocytes, pooled from multiple mice"

/dev_stage="60 day"

/lab_host="DH10B (phage-resistant)"

/clone_lib="McCarrey Eddy spermatoocytes"

/note="Organ: testis; Vector: pBluescript SK+

(Stratagene); Site 1: XhoI; Site 2: EcoRI; CDNA oligo

dt-primed [5'-(GA)10-ACTAGCTCGAGTCTTTTCTTTT-3'] and

directionally cloned using 5' linkers 5'-ATTGCGACGAG-3'

and 5'-CTCGTCCG-3'. Size selection of >400bp material

gives average insert size ranging from 1-2 kb. Library was

mass excised (from lambda-UniZAP-XR) and resulting

single-stranded phagemids were prepped and transformed

into DH10B. Library contains 98% recombinants.

References: J. Androl. 20:635-639 and Gene 25:263-269.

Library constructed and donated by J. McCarrey, Ph.D.

(Southwest Foundation for Biomedical Research, Dept. of

Genetics); excision done by E.M. Eddy, Ph.D. (National

Institutes of Health, National Institute of Environmental

Health Sciences). Original lambda-based library is

available through ATCC, catalog #63422."

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length: 597

Matches: 162

Conservative: 6

Mismatch: 15

Indels: 0

Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x CB272391 (1-597)

1 MetGlnAlaHisAlaLeuThrIleArgLeuThraSlnValGluPheProPheLeuVal 20
49 ATGAGGCTCAGCAGCTAGCTATTAGCTCACAATTAAGTAATTCCTTTTATTGTT 108
21 LeuLeuIleSerGlyValHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
109 CTTTGTATTTCTGGCGGTCTGCTGCTGTGGCATTAGCCAAAGGTGTTCCGATTTTCTGT 168
41 LeuLeuGlyValSerLeuAspIleAlaProGlyAspMetLeuAspValAlaArgArg 60
169 CTCCTTGGAGGCTTTTGACATAGCACAGGCAATGCTTGACAGGTGCGCAAGAGA 228
61 LeuSerLeuIleLeuHisProGlyCysSerThrMetSerGlyGlyValAlaIleGluHis 80
229 CTTTCTTATTCACAACTCCAGATGTTCTTACATGATGCTGAAAAAGCTATGAAACAG 288
81 LeuAlaValGlnGlyAsnArgPheHisPheAspIleLeuProPheLeuHisAlaVal 100
289 TTGGCCAAAGACGGAATTAATTCATTCTTCAATCAATCACTCAATCAATGCTAAG 348
101 AsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspValIleIleMetLeuVal 120
349 AATTGCAATTTTCTTTCACGCGACTTCACATATTAATTAAGCTAATTAACACACAG 408
121 GlnuLysGluGlyValGluValGlnIleLeuSerSerAlaAlaAspIleAlaVal 140
409 GAAAAAGAAAGAGCTTGAAGAGGCGCAATCTGTCTATCAGCTGCGACATCTCTCT 468

141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 160
469 GCGGTACAGCAATGACACAGCGTGCACCTTCGAAAGAAACACATCGGCTATTCTGTT 528
21 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyVal 180
529 TGCAGCAGAAAAATTCTCTCTCCACACTAACGACGATTAATGTTATCTGAGCTGTT 588
181 AlaSerAsn 183
589 GCAAGTAAC 597

RESULT 13

BU403563

LOCUS

60413845661 CSEQCHN59 Gallus gallus CDNA clone CHEST967014 5', mRNA

SEQUENCE

ACCESSION BU403563.1 GI:25772619

VERSION

KEYWORDS

SOURCE

ORGANISM

Gallus gallus (chicken)

Gallus gallus

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CONTACT: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 10D, UK

Tel: 0161208930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1..792

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn, Hisex"

/db_xref="taxon:9031"

/clone="CHEST967014"

/dev_stage="36"

/lab_host="DH10B"

/clone_lib="CSEQCHN59"

/note="Organ: limbs; Vector: pBluescript II KS(+); Site 1:

EcoRI; Site 2: NotI; This normalized library was

constructed from 1 million independent clones. CDNA

synthesis was initiated using an oligo(dT) primer, using

methylated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was blunted, ligated to NotI adapters, digested with

EcoRI, size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the

pBluescript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS

(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6

(1996): 791, except that a significantly longer

reannealing hybridization was used."

ORIGIN

Alignment Scores:

Pred. No.: 1,016-84

Score: 819.50

Percent Similarity: 78.99%

Best Local Similarity: 63.81%

Query Match: 59.17%

Length: 792

Matches: 164

Conservative: 39

Mismatch: 53

Indels: 3

Gaps: 1

US-10-649-273-2_copy_148_414 (1-267) x BU403563 (1-792)

```
Oy 5 AAlaLeuThrTlLeuArgLeuThrLeuValGluPheProPheLeuValLeuLeuSer 24
Db 2 GCACTTACCAATGACATGACAGACAGCAAGTGAATTCCTTCTTAGTCTTTACTCTCC 61
Oy 25 GtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 44
Db 62 GGAAGGTCATGTCAT - CTGGCAGTACGACGAGAGTTCATATTTCTTCTGCTTGACAG 120
Oy 45 SerLeuAspTlleAlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeu 64
Db 121 TCCATGATATGATGACACAGGTCGATGATGATGATGATGATGATGATGATGATG 180
Oy 65 LysHisProGluCysSerThrMetSerGlyLysAlaLleGluHisLeuAlaLysGln 84
Db 181 AAGCAACCCGAGGACGACAGCAGTGGCGGAGGAGCAATGAGCAGCTGGCTCAACC 240
Oy 85 GlyAsnArgPheHisPheAspTlleLysProProLeuHisHisAlaLysAsnGlyAspPhe 104
Db 241 GGAAGATGCGACACAGTACACTTCCATGACAGTACAGTACAGTACAGTACAGTATTT 300
Oy 105 SerPheThrGlyLeuGlnHisValThrAspLysLleLleLysLysGlyGlyGlyGly 124
Db 301 TCTTCTCCGAGCTTCAGAGCTTGTCAACAAAGCCATCTTCAGAAAGAAAGAAAGAA 360
Oy 125 GlyLleGluLysGlyGlnLleLeuSerSerAlaAlaAspTlleAlaAlaThrValGlnHis 144
Db 361 GGATATTCAGAAAGGAGAAATCTGCTCGCTTGAAGCAATCCCTGCTGTCAGACAGCAC 420
Oy 145 ThrMetAlaCysHisLeuValLysArgThrHisArgAlaLleLeuPheCysLysGlnArg 164
Db 421 GATAGTGGCTGCTCATTTATTCAGCGGACACACCGACATGCTCTTCTGCAAGAAAC 480
Oy 165 AspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyLysAlaLysAsnPhe 184
Db 481 AGCATATTTATTCACAAACCTGCACTGCTGTTATCAGAGAGAGTTCAGAGTATCAG 540
Oy 185 TyrTlleArgArgAlaLeuGlnLleLeuThrAsnAlaThrGlnCysThrLeuLeuCysPro 204
Db 541 TATATTCAGAAAGCATGCTGCTGCAATGCAACGCTTCTTCTGCTGCTGCT 600
Oy 205 ProProArgLeuCysThrAspAsnGlyLleMetLleAlaThrAsnGlyLleGluArgLeu 224
Db 601 CTTCAAGGCTGTCACCGATATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Oy 225 ArgAlaGlyLeuGlyLleLeuHisAspTlleGlyLysLysLysLysLysLysLysLys 244
Db 661 CGTGCAGAGTGTGAT - TTATACAGTACGATGATGATGATGATGATGATGATGATG 719
Oy 245 LeuGlyValAspTlleSerLysGlyValGlyLysAlaSerLleLysValPro 261
Db 720 CTGGAATG---ATATTCACAAAGAGTGAAGAGATCCATCAAGTGCCA 767

RESULT 14
CN823245 730 bp mRNA linear EST 02-JUN-2004
LOCUS CN823245
DEFINITION Oa_splbn_04N08_M13reverse Sheep spleen/brain pSPORT1 library Ovis
ACCESSION CN823245
VERSION CN823245.1 GI:47951314
KEYWORDS EST.
SOURCE Ovis aries (sheep)
ORGANISM Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Caprinae; Ovis.
1 (bases 1 to 730)
AUTHORS Gossner, A. and Hopkins, J.
TITLE Ovine spleen/brain cDNA library
JOURNAL Unpublished (2004)
COMMENT Contact: J Hopkins
```

Veterinary Biomedical Sciences
University of Edinburgh
Summerhall Square, Edinburgh, EH9 1QH.
Email: j.hopkins@ed.ac.uk
Plate: 04 row: N column: 08
Seq primer: M13reverse
High quality sequence start: 6
High quality sequence stop: 550.
Location/Qualifiers
1..730
/organism="Ovis aries"
/mol_type="mRNA"
/db_xref="taxon:9940"
/clone="Oa_splbn_04N08"
/clone_1lb="Sheep spleen/brain pSPORT1 library"
/note="Vector: pSPORT1"

ORIGIN

Alignment Scores:
Pred. No.: 2.3e-84 Length: 730
Score: 816.00 Matches: 157
Percent Similarity: 93.30% Conservative: 10.
Best Local Similarity: 87.71% Mismatches: 12
Query Match: 58.92% Indels: 0
DB: 7 Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x CN823245 (1-730)

```
Oy 1 MetGluAlaHisAlaLeuThrTlleArgLeuThrAsnLysValGluPheProPheLeuVal 20
Db 193 ATGAGAGCTCATGACATCTACTATTAGGTTAACAAAATGAGTGAATTTCCATTTTATGTT 252
Oy 21 LeuLeuLleSerGlyGlyHisCysLeuLeuAlaLeuValGlnGlyValSerAspPheLeu 40
Db 253 CTTTGAATTTCTGGAGTCTCATGTCCTTTGGCATTACTGAGGAGTTTCAGATTTCTT 312
Oy 41 LeuLeuGlyLysSerLeuAspTlleAlaProGlyAspMetLeuAspLysValAlaArgArg 60
Db 313 CTTCTTGGAAGCTTTTGACATAGCACACGCGACATGCTTGACAGTACAGTACAGAAAGA 372
Oy 61 LeuSerLeuLleLysHisProGluCysSerThrMetSerGlyLysAlaLleGluHis 80
Db 373 CTTTGTCTTAATTAACATCCAGAGTCTCCACCATGAGTGGGAGGCTATGAAACAT 432
Oy 81 LeuAlaLysGlnGlyAsnArgPheHisPheAspTlleLysProProLeuHisHisAlaLys 100
Db 433 TTGGCCAAACAGGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 492
Oy 101 AsnGlyAspPheSerPheThrGlyLeuGlnHisValThrAspLysLleLleMetLysLys 120
Db 493 AATTGTAATTTTCTTTTCTGACATTCATGATGATGATGATGATGATGATGATGATG 552
Oy 121 GlnLysGlnGlnGlyLleGlyGlyGlyGlnLleLeuSerSerAlaAlaAspTlleAlaAla 140
Db 553 GAAABACAGAGAGGTATCGACAGGCGGACATCTGCTTCAGTCGACATGCTGCT 612
Oy 141 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaLleLeuPhe 160
Db 613 GCACTTCAGACACACCGTGGCTGCGACATTCGCAAAAGAACACATCGGCTATTTCTG 672
Oy 161 CysLysGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGly 179
Db 673 TGCAGAGAGAGGCTGTTATCATCAAGTACGACAGTACTGTTGATCTGAGAGC 729

RESULT 15
AV602901 545 bp mRNA linear EST 27-NOV-2001
LOCUS AV602901/c
DEFINITION AV602901 Bos taurus kidney fetus Bos taurus cDNA clone B1K1013A07
ACCESSION AV602901
VERSION AV602901.1 GI:9725227
KEYWORDS EST.
SOURCE Bos taurus (cow)
```

ORGANISM

Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.

REFERENCE
AUTHORS

1 (bases 1 to 545)
Takesuga, A., Hirotsune, S., Itoh, R., Jirohono, A., Suzuki, H., Aso, H.
and Sugimoto, Y.

TITLE

Establishment of a high throughput EST sequencing system using
poly(A) tail-removed cDNA libraries and determination of 36,000
bovine ESTs

JOURNAL
MEDLINE
PUBMED

Nucleic Acids Res. 29 (22), E108 (2001)
21570554
11713328

COMMENT

Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-Shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazunugi@cocoa.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
Location/Qualifiers

FEATURES
source

1..545
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="EIKI013A07"
/tissue_type="kidney"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_lib="Bos taurus kidney fetus"
/note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"

ORIGIN

Alignment Scores:

pred. No.: 2,14e-81 Length: 545
Score: 789.00 Matches: 147
Percent Similarity: 90.11% Conservative: 17
Best Local Similarity: 80.77% Mismatches: 17
Query Match: 56.97% Indels: 1
DB: 1 Gaps: 0

US-10-649-273-2_COPY_148_414 (1-267) x AV602901 (1-545)

QY 85 G|VAsnArpPheHisPheAspIleLeuProPheLeuHisHis-AlaIysAsnCyAspPh 104
DB 544 GGAATATGATTCATTTTGAATTTCCANCTCCATGCACGCTCTAAATTTGATTT 485
QY 104 eSerPheThG|YleuGlnHisValThrAspLysIleIleMetLysLysGluGlu 124
DB 484 TTCTTTTTCGACTTCACACGTTATGATPAAGATGATATGCAAAAGAAANAGA 425
QY 124 uGlyIleGluLysGlyGlnIleLeuSerSerAlaAlaAspIleAlaThrValGlnH 144
DB 424 AGGTATCGACAGGGGCGCTCTTCAGCTCGCGACATTCGCTGTGCGTCCAGCA 365
QY 144 sThrMetAlaCyHisIleuValLysArgThrHisArgAlaIleLeuPheCyLysGlnAr 164
DB 364 CACCGTGGCTGCGACATGCAAAAGAACACATGCTCTTCGTCTTCGACAGCAG 305
QY 164 GAAPLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPh 184
DB 304 AGGCTTCTTACATCAGAGTAAGCAGTACTGTTGATCTGAGGCGTCGCAAGTAACTT 245
QY 184 eTyrlLeaArgArgAlaLeuGluIleLeuThrAsnAlaThrGlnCyThrLeuLeuCyPr 204
DB 244 ATATATCCGAAAGCCCTGGAATTTGACCAATGCAACACAGTCACTTTGCTGTGCC 185
QY 204 oPrPcArArgLeuCyThrAspAsnGlyIleMetIleAlaThrAsnGlyIleGluArgLe 224
DB 184 GCCCCCACTGCACTGACATGCAACGCGTATGATTCATGCAATGATGTTGAAAGACT 125

QY 224 uArgAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyGluProLysCyPr 244
DB 124 ACGTGTGGCTTGGGCAATTTTACACAAACAGGAGCATCCGCTACGAACCAAAATGCC 65
QY 244 oLeuGlyValAspLysSerLysGluValGlyGluAlaSerIleLysValProGlnLeu 264
DB 64 TCTTGGAGTATGATATATCCAAAGAAAGTTGAGAAAGCTGCTATMAAAGTCCAAAGTTAA 5
QY 264 sMet 265
DB 4 AATG 1

Search completed: June 17, 2005, 01:30:00
Job time : 2296.07 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 16, 2005, 17:01:19 ; Search time 2048.13 Seconds

(without alignments)
4441.797 Million cell updates/sec

Title: US-10-649-273-2_COPY_176_414
Perfect score: 1240
Sequence: 1 ILALVQGVSDPFLGKSLDI.....DISKEVGEASIKVPQKMEI 239

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+g2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10649273/runat_15062005_111416_6043/app.query.fasta_1.1429
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=-1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10649273@cgn2.1 8076@runat_15062005_111416_6043 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELCK=100 -XLONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gsa1:
9: gb_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1118	90.2	870	5	BQ423651 AGENCOURT
2	1057	88.5	840	5	BQ636028
3	1079	87.0	2284	3	AK045669 Mus muscu
4	1071	86.4	1622	3	AK011265 Mus muscu
5	930	75.0	852	5	BX391919
6	526	74.7	1609	3	BCU30671
7	908	73.2	701	2	BC030611 Mus muscu
8	876	70.6	658	7	BE740611
9	838	67.6	637	7	CP362328
					CK941819 4065407 B

10	799	64.4	922	5	BQ61028	BQ61028 AGENCOURT
11	789	63.6	345	1	AV602901	AV602901
12	784	63.2	490	6	CB852881	CB852881 UI-CF-FN0
13	783	63.1	822	7	CF257246	CF257246 pha008_g0
14	753	60.7	866	5	BU127463	BU127463 603114407
15	749	60.4	736	7	CK365185	CK365185 AGENCOURT
16	746	60.2	723	5	BU61251	BU61251 603502215
17	731.5	59.0	792	5	BU403563	BU403563 604138456
18	717	57.8	634	2	AM601179	AM601179 RCL-BT025
19	708	57.1	736	6	CA057753	CA057753 ssa1rpb54
20	707.5	57.1	701	5	BU621780	BU621780 UI-H-FL1-
21	698	56.3	696	5	BU302606	BU302606 603739448
22	685	55.2	597	6	CB272391	CB272391 ma157912
23	683	55.1	909	5	BX756548	BX756548 BX756548
24	682	55.0	548	7	CO880741	CO880741 Boycen_09
25	671	54.1	682	7	BB043703	BB043703 BB043703
26	671	54.1	730	7	CN823245	CN823245 Oa.sp1bn
27	668	53.9	706	5	BU202465	BU202465 603849052
28	666	53.7	869	5	BX754527	BX754527 BX754527
29	647	52.2	919	7	CF407294	CF407294 CH3#047_E
30	646	52.1	484	1	AJ670918	AJ670918 AJ670918
31	645	52.1	878	5	BD508917	BD508917 CDA93-E05
32	642	51.8	1173	6	CD508917	CD508917 CDA93-E05
33	638	51.5	1082	5	BX359023	BX359023 BX359023
34	625	50.4	506	2	BF415802	BF415802 UI-R-CA1-
35	614	49.5	861	5	BU246158	BU246158 603779906
36	611	49.3	1171	5	BU261605	BU261605 603501763
37	604	48.7	424	1	AA273921	AA273921 VB99C03.r
38	600	48.4	863	5	BU376295	BU376295 603808890
39	594	47.9	389	4	BM744822	BM744822 K-EST0018
40	591	47.7	812	5	BU246489	BU246489 603784202
41	589	47.5	357	5	BQ672554	BQ672554 AGENCOURT
42	581	46.9	704	5	BU327284	BU327284 603491570
43	576	46.5	825	7	CR444994	CR444994 CR444994
44	573	46.2	385	1	AJ647827	AJ647827 AJ647827
45	572.5	46.2	613	1	AA920105	AA920105 v225h05.r

ALIGNMENTS

RESULT 1
BQ423651
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITL
JOURNAL
COMMENT

BQ423651 870 bp mRNA linear EST 23-MAY-2002
AGENCOURT 7790948 NIH_MGC_72 Homo sapiens CDNA clone IMAGE:6065828
5', mRNA sequence.
BQ423651
BQ423651.1 GI:21118966
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 870)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: ATCC/CDT/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LHAM13342 row: 1 column: 21
High quality sequence stop: 710.
Location/Qualifiers
1. 870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6065828"

FEATURES

source

/issue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NH1 MGC 72"
 /note="Organ: skin; Vector: PCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 36-119 Length: 870
 Score: 1118.00 Matches: 220
 Percent Similarity: 98.22% Conservative: 1
 Best Local Similarity: 97.78% Mismatches: 2
 Query Match: 90.16% Indels: 2
 DB: 5 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x BQ423651 (1-870)

QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
 DB 196 CTTTGGCATTACTTCAAGAGATTTCAGATTTCCTGCTTGGAAAGCTTTGGACATA 255
 QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysSHISProGlu 40
 DB 256 GCACCAAGAGACATGCTTGCACAGTGCGCAAGAACTTTCTTTAATPAAACATCCAGAG 315
 QY 41 CysSerThrMetSerGlyGlyLysAlaIleGlnHisLeuAlaLysGlnGlyAsnArgPhe 60
 DB 316 TCTCTCCACATGAGTGCGTGGGAAAGCCATAGAACATTTGGCCAAACAGAAATAGATTT 375
 QY 61 HisPheAspIleLysProPheLeuHisSHISAlaLysAsnCysAspPheSerPheThrGly 80
 DB 376 CATTGTAACATCAACCTCCCTTGCATCATCTPAAATATGATTTTCTTTTACTGGA 435
 QY 81 LeuGlnHisValThrAspLysIleIleMetLysGlyLysGlnGlyIleGluLys 100
 DB 436 CTTCAACACACTTACTGATAAATATATGAAAGAAAGGAAAGGATTTGAGAGAG 495
 QY 101 GlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCys 120
 DB 496 GGGCAAAATCTGCTCTTCAGCAGCAGACATTTGCGCCACAGTACAGCACAAATGGCATGT 555
 QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuPro 140
 DB 556 CATCTTGTAAGAAAGAACATCGGGCTATTCTGTTTGAAGACAGAGACTGTACTCT 615
 QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaLysSerAspPheThrIleArgArg 160
 DB 616 CAAAATATGCACTACTGCTGTGATCTGCTGCTGCCAAGTAATTTCTATATCCGAGA 675
 QY 161 AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
 DB 676 GCTCTGGAATTTTAAACAAACGACACAGTGCATTTGGTGGTCTCTCCACAGACTA 735
 QY 181 CysThrAspAsnGlyLysMetIleAlaThrAspGlyIleGlnArgLeuAlaGlyLeu 200
 DB 736 TGCACATGATTAATGCACTTATGATTCATGAAATGATTAATTAAGACATCACTGCGCTTG 795
 QY 201 GlyIleLeuHisAspIleLeuGlyLysArgTyGlnProLysCysSerProLeuGlyValAla 220
 DB 796 GGCATTTTATCATACATTAAGAGCATCCGCTATGAAACCAAAATGTCCTCTTGGAGTAG 855
 QY 220 spIleSerLys 223
 DB 856 ACATATCAAAA 866
 RESULT 2
 BQ636028 640 bp mRNA linear EST 15-JUN-2002
 LOCUS hd03d11.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
 DEFINITION Homo sapiens cDNA clone hd03d11 5', mRNA sequence.
 ACCESSION BQ636028

VERSION BQ636028.1 GI:21760487

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS Rukayate; Metaxoa; Chordata; Craniata; Vertebrata; Euteleostomi;

TITLE 1 (bases 1 to 640)

Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Bahal,A.,

Touchman,J.W., Bouffard,G., Smith,D., and Peterson,K.

Expressed sequence tag analysis of human retina for the NEIBank

Project: Retbinding, an abundant, novel retinal cDNA and alternative

splicing of other retina-preferred gene transcripts

Mol. Vis. 8 (4), 196-204 (2002)

JOURNAL 22103461

MEDLINE 12107411

PUBMED

COMMENT

Contact: Wistow G

Section on Molecular Structure and Function

National Eye Institute

6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 03 row: d column: 11

Seq primer: M13RPI reverse primer (ABI).

Location/Qualifiers

1..640

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="hd03d11"

/issue_type="Retina"

/dev_stage="Adult"

/lab_host="EMDHD10B"

/clone_lib="Human Retina cDNA (Un-normalized,

unamplified): hd/he"

/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue

was dissected from two 80 year old donors with no observed

eye disease. 100ug of total RNA was used for library

construction. A directionally cloned cDNA library in the

pSPORT vector (Life Technologies) was constructed at

Bioserve Biotechnology (Laurel MD) essentially following

the protocols of the Superscript Plasmid System full

details of which are contained in the manufacturer's

instruction manual (http://www.lifetech.com/). First

strand synthesis was carried out using a Not I

primer-adaptor

15'-pGACTGATCTTACATCGCGAGCGCGCC(T)15-3'. EST analysis

was performed on the unamplified library at the NIH

Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:

Pred. No.: 5,416-117 Length: 640
 Score: 1097.00 Matches: 209
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 88.47% Indels: 0
 DB: 5 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x BQ636028 (1-640)

QY 31 ArgArgLeuSerLeuIleLysPheProGluCysSerThrMetSerGlyGlyValAlaIle 50
 DB 3 AGAAGACTTTCTTTAATPAAACATCCAGAGCTCCACCATGATGTGGGAAAGCCACTA 62
 QY 51 GlnHisLeuAlaLysGlnGlyAsnArgPheHisAspIleLysProProLeuHisSHIS 70
 DB 63 GAACATTTGGCCCAACAGAAATAGATTTCATTTTACATCAAACTCCCTTGCAATAT 122
 QY 71 AlaLysAsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMet 90
 DB 123 GCTAAATATGTGATTTTCTTTTACTGCACTTCAACCGTTACTGATTAATATATATG 182

OY 91 LysylSGlyLeuGlyGluGlyGlyIleGlyLeuGlyGlyIleLeuSerSerAlaAlaAspIle 110
 DB 183 AAAAAAGAAAAAGAGAGATATGAGAGAGGCAATCTGCTTCAGACAGACAT 242
 OY 111 AlaAlaThrValGlnHisThrMetAlaCysHisLeuValIleValArgThrHisArgAlaIle 130
 DB 243 GGTGGCAGAGTACAGACACAAATGAGCATCTGTGTGAAAGAACACATCGGCTATT 302
 OY 131 LeuPheCysLeuGlnArgAspLeuLeuProGlnHisAsnAlaValLeuValAlaSerGly 150
 DB 303 CTGTTTGTAGACAGAGAGACCTGTTTACCTCAAAATATATGACATCTGCTTCATCTGCT 362
 OY 151 GlyValAlaSerAsnPheTyrIleArgArgAlaLeuGluIleLeuThrAsnAlaThrGln 170
 DB 363 GGTGTGCAAGTAACTTATATATCCGAGAGCTCTGGAATTTTAAACAAACGACACAG 422
 OY 171 CysThrLeuLeuCysProProPheProAlaGluCysThrAspAsnGlyIleMetIleAlaTTP 190
 DB 423 TCCACTTGTGTGTCTCTCTCCAGACTATGACATGATATGATGATTAATGATTCATGG 482
 OY 191 AsnGlyIleGluArgLeuArgAlaGlyLeuGlyIleLeuHisAspIleGlyGlyIleArg 210
 DB 483 AATGGATTGAAAGACTACGCTGCTGGCTTTTACATGACATGAAAGGATCCGC 542
 OY 211 TyrGluProLysCysProLeuGlyValAspIleSerIleGlyValAlaSerIle 230
 DB 543 TATGAACCAAAATGCTCTCTTGAAGTACATATCAAAAGAGTTGAGAGAGCTTCATA 602
 OY 231 LysValProGlnLeuLysMetGluIle 239
 DB 603 AAGGTACCAATTAATAATGAGATA 629
 RESULT 3
 LOCUS AK045669 2284 bp mRNA linear HTC 03-APR-2004
 DEFINITION Mus musculus adult male corpora quadrigemina cDNA, RIKEN
 full-length enriched library, clone:B230219017 product:similar to
 PUTATIVE SIALOGLYCOPROTEASE TYPE 2 (Homo sapiens), full insert
 sequence.
 ACCESSION AK045669
 VERSION AK045669.1 GI:26337528
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Komno, H., Akiyama, J., Nishi, K., Kiteunui, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
 Yonekura, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 JOURNAL sequencing pipeline with 384 multicapillary sequencer
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE
 AUTHORS 4
 TITLE The RIKEN Genome Exploration Research Group Phase II Team and the
 JOURNAL FANTOM Consortium.
 REFERENCE Functional annotation of a full-length mouse cDNA collection
 AUTHORS Nature 409, 685-690 (2001)
 TITLE 5
 JOURNAL The FANTOM Consortium and the RIKEN Genome Exploration Research
 REFERENCE Group Phase I & II Team.
 ANALYSIS Analysis of the mouse transcriptome based on functional annotation
 OF 60,770 full-length cDNAs
 NATURE 420, 563-573 (2002)
 TITLE 6 (bases 1 to 2284)
 JOURNAL Adachi, J., Aizawa, K., Akiyama, T., Arakawa, T., Bono, H., Carninci, P.,
 REFERENCE Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 AUTHORS Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
 Horii, F., Imoto, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komno, H., Kouda, M.,
 Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeuchi, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 PHYSICAL AND CHEMICAL RESEARCH (RIKEN), Laboratory for Genome
 EXPLORATION RESEARCH GROUP, RIKEN GENOMIC SCIENCES GROUP (GSC),
 KANAGAWA 230-0045, Japan (E-mail: genome-res@gs.riken.jp,
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 COMMENT cDNA library was prepared and sequenced in Mouse Genome
 ENCycloPedia Project of Genome Exploration Research Group in Riken
 GENOMIC SCIENCES Center and Genome Science Laboratory in RIKEN.
 DIVISION Division of Experimental Animal Research in Riken contributed to
 PREPARE mouse tissues.
 PLEASE visit our web site for further details.
 URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/.
 FEATURES
 SOURCE Location/Qualifiers
 1..2284
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM,DB:B230219017"
 /db_xref="taxon:10090"
 /clone="B230219017"
 /sex="male"
 /tissue_type="corpora quadrigemina"
 /clone_id="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 221..1465
 /note="unnamed protein product; putative
 similar to PUTATIVE SIALOGLYCOPROTEASE TYPE 2 [Homo
 sapiens] (SPTR|Q9H4B0, evidence: PASTY, 80%ID, 100%length,
 match=142)"
 /codon_start=1
 /protein_id="BAC32450.1"
 /db_xref="GI:26337529"
 /translation="MIMRRRTAGALIPRKSKVYGLRPSVHPRTLSCHLVLEIET
 SCDDTGAAYVDENGVAGLALHSTQVHLKGTGVIYPPAQQOLHRENTIORIYEETLSAC
 RITPDSLSAIVTTRKGLASLVGGLSLDVLVDFKPFPIHNEHAALITRLTK
 VEPFLVNLISGCHLVLVGVDFLGLSLDLPAMDVLRSLIKHPGST
 MSGKALBOALDOGRFPHTINPQWAKNCDPSFGQIHITDKLITKEKEGIEKG
 OLSGKALIAAVOAHATFCHLAKTRHAILCCKQNLSPAAVLVSGGVASNYIR
 KALIVANATCTILCPPTCTDNGIMIANNGIERLRLAGLGVLDVDIRBPCPL
 GVDISREVAEAAIVPRKML"
 2262..2267
 /note="putative"
 2284
 /note="putative"
 ORIGIN
 polyA_signal
 polyA_site


```
source
1..1622
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:2610001M19"
/db_xref="taxon:10090"
/clone="2610001M19"
/issue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days embryo"
207..1451
/note="unnamed protein product; putative
similar to PUTATIVE SIALOGLYCOPROTEASE TYPE 2 [Homo
sapiens] (SPRR|Q9H4B0, evidence: FASTY, 80%ID, 100%length,
match=1242)"
/protein_id="BA27506.1"
/db_xref="GI:12847276"
/translation="MLMRLTGAIPKPKSKYVGLRFRSVHPTLSCHKLVGIET
SCDDTGAADVDEGTGVALHSGVLSQVHLTGQIVPPVAQQLRENIQRIVEETLSAC
RTSPDSIAITIKRGLALSGVLSFSLVNOFKKPIPIHMEAHALITRLTK
VPPRLVILISGHCILAVGVSPDLISGLSDIAPGMDLDRVARRLSIKPECT
MSGKAIKQAKGNRPHITINPPMNAKCDSPFGLOHTTKLITKKEGIEKE
QILSSADIAAVQAHATACHLAKTRHAILFCQKQLSPANVLVSGVANSLYR
KALEIVANAQCTLCPPRLCTDNGIMIAMNGIERLRAGLGLHVEDIRYKPCPL
GVDISREVAEAAIKVPRLMKWL"
1605..1610
/note="putative"
1622
/note="putative"
1622
/note="putative"

polya_signal
polya_site

ALIGNMENT SCORES:
Pred. No.: 2.16e-113 Length: 1622
Score: 1071.00 Matches: 204
Percent Similarity: 91.21% Conservative: 14
Best Local Similarity: 85.36% Mismatches: 21
Query Match: 86.37% Indels: 0
DB: Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x AK01265 (1-1622)
QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
DB 732 CTTTGGCATTAGTCCAGAGGTTGTCGATTTCTGCTCTTGGCAAGCTTTTGGACATA 791
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 792 GCAACGAGGACATGCTTGACAGGTGGCAAGAAACATTTCTTAAATCAACATCCAGAA 851
QY 41 CysSerThrMetSerGlyLysValAlaIleGluHisLeuAlaLysGlnGlyAsnArgPhe 60
DB 852 TGTTCACATAGAGTGTGAGAAAGCTATAGAACAGTTGGCCAAAGACGGAATAGATTC 911
QY 61 HisPheAspIleLysProGluLeuHisAlaLysAsnCysAspPheSerPheThrGly 80
DB 912 CATTTAATCAATCAATCCATGAGAAATGCTAAGAAATGCGATTTTCTTACCGGGA 971
QY 81 LeuGlnHisValThrAspLysIleIleMetLysGlyLysGlnGlyLysGlnGlyLeu 100
DB 972 CTTCAACATTTCTGATAGCTAATTAACACACAGAAAGAAAGAAAGGCAATGGAGAG 1031
QY 101 GlnGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCys 120
DB 1032 GAGCAATCTCGTCATAGCTGAGACATTTGCTGCGGTGACGATGCAACGCGTGC 1091
QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuPro 140
DB 1092 CACCTTGCGAAAGAACACATCCGCGTATCTGTTTGGCAAGCAAAATTTCTCTCT 1151
QY 141 GlnAsnAspAlaValLeuValAlaSerGlyValAlaLysSerAsnPheThrIleArgArg 160
DB 1152 CCAGTAAACGACAGTATTGATTGATTCGAGGTTGCAAGTAACTTGTAATACCGAANA 1211
```

```
QY 161 AlaLeuGluIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProProArgLeu 180
DB 1212 GCATTGAAATTTGCGCAAAATGCAACGACGACGTTGTTGTGTCTCCACCTCCAGACCTG 1271
QY 181 CysThrAspAsnGlyIleMetIleAlaThrAsnGlyIleGluArgLeuArgAlaGlyLeu 200
DB 1272 TGCACCTACCAATGGCATCATGATTTGCAATGGAATTTGAAAGATTACGTTGCGCTTG 1331
QY 201 GlyIleLeuHisAspIleGluGlyIleArgGlyArgProLysCysProLeuGlyValAsp 220
DB 1332 GCGCTTTACATGATGTAGGAAGACATCCGATATGACCAAAATGCTCTTGAGTAAAC 1391
QY 221 IleSerLysGlyValGlyValAlaSerIleLysValProGlnLeuLysMetGluIle 239
DB 1392 ATATCCAGAGAAATTGCAAGAGCTGCATTAAGATGACGCGATTAAATATGCACTT 1448

RESULT 5
BX391919/c 852 bp mRNA linear EST 28-APR-2004
LOCUS BX391919 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DK001YE02 3-PRIME, mRNA sequence.
ACCESSION BX391919
VERSION BX391919.2 GI:46846154
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 852)
AUTHORS Li,W.-B., Gruber,C., Jessee,J. and Polayes,D.
TITLES Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced gi:30611736.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by life technologies, a
division of invitrogen. This sequence belongs to sequence cluster
1240.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BA1035ZE07_CS03317_1&c=1240.r

FEATURES
source
1..852
/location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YE02"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dt)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized."

ALIGNMENT SCORES:
Pred. No.: 2.28e-97 Length: 852
Score: 930.00 Matches: 180
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 75.00% Indels: 0
DB: Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x BX391919 (1-852)
QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
```

Db 542 CTTGTTGGCAATTAGTAAAGAGATTTCAGATTTCCTCTTCTTGGAAAGCTTTGGACATA 483
Oy 21 AAlAProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
Db 482 GCACCAAGGACGACATGCTTCAACAGGTGGCAAGAGACTTTCTTAATCAACATCCAGAG 423
Oy 41 CysSerTherMetSerGlyGlyLysAlaIleGlnHisLeuAlaLysGlnGlyAsnArgPhe 60
Db 422 TGCCTCCACCATGATGCTGGGAAAGCCATAGAACATTGGCCAAACAGAAATAGATT 363
Oy 61 HisPheAspIleLysProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGly 80
Db 362 CATTTTGACATCAAACTCCCTGCAATCAAGCTAAATAATGATATTTCTTTTACCGGA 303
Oy 81 LeuGlnHisValThrAspLysIleIleMetLysLysGlnLysGlnGlnGlyIleGluLys 100
Db 302 CTTCAACAGCTTACTATTAATTAATATATCAAAAAGAAAAGAGAGATTTAGAGAG 243
Oy 101 GylGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCys 120
Db 242 GGGCAAAATCTGTCCTTCAAGAGACATTTGCTGCCACACTACAGACACAAATGCGATGT 183
Oy 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
Db 182 CATCTTGAAAAAGAACACATCGGCTATTTCTGTTTGAAGACAGAGACTTTTACT 123
Oy 141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnProPheTrpLeuArg 160
Db 122 CAAAATAAGTCAGATCTGCTTGCATCTGGTGTGTCGCAAGTAATCTTAATATCCGAGA 63
Oy 161 AAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
Db 62 GCTCTGGAATTTTACCAACAGCAACAGCTGCACTTTGTGTGTCTCTCTCCACACTA 3

RESULT 6
BC030671 1609 bp mRNA linear HTC 19-NOV-2003
LOCUS Mus musculus O-sialoglycoprotein endopeptidase-like 1, mRNA (cDNA
DEFINITION clone IMAGE:1226118), containing frame-shift errors.
ACCESSION BC030671
VERSION BC030671.1 GI:21040459
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1609)
AUTHORS Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.L., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.G., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.B., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carinci, P., Prange, C., Raha, S.S., Loguercio, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McSwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, Y.K., S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schultz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Scherer, A., Schein, J.E., Jones, S.J., Jones, S.J., and Marra, M.A.
human and mouse cDNA sequences of more than 15,000 full-length
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 1609)
AUTHORS Strausberg, R.

TITLE Direct Submission
JOURNAL Submitted (20-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Marcello Bento Soares, Ph.D.
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
contact: amadad@systemsbiology.org
Anup Madan, Jessica Fahy, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

FEATURES
source
1. 1609
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1226118"
/tissue_type="Thymus gland, mouse"
/clone_id="Soares_thymus_2NBMT"
/lab_host="DH10B"
/note="Vector: pT73-Pac"

ORIGIN
Alignment Scores:
Pred. No.: 1,656-96 Length: 1609
Score: 926.00 Matches: 183
Percent Similarity: 81.59% Conservative: 12
Best Local Similarity: 76.57% Mismatches: 20
Query Match: 74.68% Indels: 24
DB: 3 Gaps: 1

US-10-649-273-2_copy_176_414 (1-239) x BC030671 (1-1609)
Oy 1 LeuValAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
Db 749 CTTGTTGGCAATTAGTCAAGGATGTTCCGATTTCTGCTCTTGGGAAGTCTTTGGACATA 808
Oy 21 AAlAProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
Db 809 GCACCAAGGACGACATGCTTCAACAGGTGGCAAGAGACTTTCTTAATCAACATCCAGAA 868
Oy 41 CysSerTherMetSerGlyGlyLysAlaIleGlnHisLeuAlaLysGlnGlyAsnArgPhe 60
Db 422 TGCCTCCACCATGATGCTGGGAAAGCCATAGAACATTGGCCAAACAGAAATAGATT 928
Oy 61 HisPheAspIleLysProProLeuHisHisAlaLysAsnCysAspPheSerPheThrGly 80
Db 362 CATTTTGACATCAAACTCCCTGCAATCAAGCTAAATAATGATATTTCTTTTACCGGA 988
Oy 81 LeuGlnHisValThrAspLysIleIleMetLysLysGlnLysGlnGlnGlyIleGluLys 100
Db 302 CTTCAACAGCTTACTATTAATTAATATATCAAAAAGAAAAGAGAGATTTAGAGAG 1108
Oy 101 GylGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCys 120
Db 242 GGGCAAAATCTGTCCTTCAAGAGACATTTGCTGCCACACTACAGACACAAATGCGATGT 1168
Oy 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
Db 182 CATCTTGAAAAAGAACACATCGGCTATTTCTGTTTGAAGACAGAGACTTTTACT 1168

QY 141 GlnAsnAsnAlaValLeuValAlaSerGlyValAlaSerAspHeterYrileargArg 160
DB 1169 CCAGCTAACGCAATATTAGTTATCTGAGAGGCTTGCAGTAATCTGATCATCCGAAA 1228
QY 161 AlaLeuGluIleLeuThraAsnAlaThrGlnCysThreLeuGlyProProPheArgLeu 180
DB 1229 GCATTGGAAATGTTCCCAATGCAACGAGTGCACGTTGTGTGTCCACCTCCCAAGACTG 1288
QY 181 CysThraAsnGlyIleLeuIleAlaTrpAsnGlyIleGluArgLeuAlaGlyLeu 200
DB 1289 TCCACGACAAATGCAATGCAATGATTCGA----- 1315
QY 201 GlyIleLeuHisAspIleGluGlyIleArgGlyLeuProPheCysProLeuGlyValAsp 220
DB 1316 -----TGAATCTCTCTTGCAGTAC 1336
QY 221 IleserLeuGlyValGlyGluAlaSerIleLeuValProGlnLeuGlyMetGluIle 239
DB 1337 AATTCAGAGAGAGTGCAGAAAGCTGCATTAAGTACCGGATTTAAATGCACTT 1393
RESULT 7
BE740611 701 bp mRNA linear EST 15-SEP-2000
LOCUS 601595739F1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3949640 5',
DEFINITION mRNA sequence.
ACCESSION BE740611 GI:10154603
VERSION BE740611.1
KEYWORDS Homo sapiens (human)
SOURCE EST.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMLU)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMLU at: image.llnl.gov
Plate: LHC9814 row: n column: 09
High quality sequence stop: 701.
Location/Qualifiers
1..701
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3949640"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_9"
/note="Organ: ovary; Vector: pOT87; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
Alignment Scores:
Pred. No.: 6,31e-95 Length: 701
Score: 908.00 Matches: 178
Percent Similarity: 98.91% Conservative: 3
Best Local Similarity: 97.27% Mismatches: 0
Query Match: 73.23% Indels: 2
DB: 2 Gaps: 0

US-10-649-273-2_copy_176_414 (1-239) x BE740611 (1-701)
QY 1 LeuLeuAlaLeuValGlnGlyValSerAspHeterLeuLeuGlyLysSerLeuAspIle 20
DB 155 CTGTGGCATTAAGTTCAAGAGGTTTCAGATTTCCTCTTGGAAATCTTTGGACATA 214
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 215 GCACAGAGTGACATGCTTGAAGAGTGCAAGAGACTTCTTTAATAAATCATCCAGAG 274
QY 41 CysSerThrMetSerGlyGlyValAlaIleGluHisLeuAlaLysGlnGlyAsnArgPhe 60
DB 275 TGCTCCACCAAGAGAGTGAGGAGAAAGCCATTAACATTTGGCCAAACAGAAATGATTT 334
QY 61 HisPheAspIleLysProProLeuHisAlaLysAsnGlyAspPheSerPheThrGly 80
DB 335 CATTTTGACATCAACCTCTCTGATCATCTCAATAAATGTGATTTCTTTTACTGCA 394
QY 81 LeuGlnHisValThrAspLysIleIleMetLys--LysGluLysGluGluGlyIleGlu 100
DB 395 CTTCAACAGCTTACTGATTAATAATATATGAAAACAGAAACAGAGAGATTTGAGA 454
QY 100 YsgIyGlnIleLeuSerSerAlaAlaAspIleAlaIleValGlnHisThrMetAlaC 120
DB 455 AGGGGCAAAATCTGCTTCAACAGACAGACATTTGCTGCACAGTACAGACCAATGGCAT 514
QY 120 YsgIleLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuP 140
DB 515 GTCAATCTTGTAAGAAACACATCGGGGTATTTGTTTGAAGAGAGACTTGTTAC 574
QY 140 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAspHeterYrileargArg 160
DB 575 CTGAATTAATGCACTGCTGTTGATCTGTGTGTGTCGCAAGTACTTCTATATCCGA 634
QY 160 GlnAlaLeuGluIleLeuThraAsnAlaThrGlnCysThreLeuGlyProProPheArg 180
DB 635 GAGCTTCGAAATTTTAACAACGACACAGTGCATTTGTGTGTCTCTCCCAAGC 694
QY 180 euCys 181
DB 695 TATGC 699
RESULT 8
CF362328 658 bp mRNA linear EST 25-AUG-2003
LOCUS CF362328 MARC 3P1G Sus scrofa CDNA 3', mRNA sequence.
DEFINITION CF362328
ACCESSION CF362328.1 GI:34161882
VERSION EST.
KEYWORDS Sus scrofa (pig)
SOURCE Sus scrofa
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 658)
AUTHORS Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Fox,J.,
Wise,T.A., Nommaman,D.J., Wray,J.E. and Keele,J.W.
TITLE A second set of porcine ESTs from a pooled-tissue normalized
library
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TP
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: SRG8015 row: H column: 12
Seq primer: TAGAAGCAGAGTCAGG.
FEATURES
source
1..658
Location/Qualifiers

/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 3Pig"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including brain, liver, muscle, placenta/endometrium,
ovary, testes, and bone marrow."

ORIGIN

Alignment Scores:
Pred. No.: 3,07e-91 Length: 658
Score: 876.00 Matches: 165
Percent Similarity: 93.68% Conservative: 13
Best Local Similarity: 86.84% Mismatches: 12
Query Match: 70.65% Indels: 0
DB: 7 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x CF362328 (1-658)

```

OY      1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlnGlySerLeuAspIle 20
DB      630 CTTTGGCATTGATTGAGAGAGATTTCAGATTTCCTCTTGACAGCTTTGGACATA 571
OY      21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB      570 GCACCAAGTGCAGTCCTTACACAGAGTACAGAGACCTTTCTTAATTAACATCCAGAG 511
OY      41 CysSerThrMetSerGlyGlyLysAlaIleGlnHisLeuAlaLysGlnGlyAsnArgPhe 60
DB      510 TGCTCCACCATGATGCTGTGGAGAGGCCATTAAGACATTTTCCAAACAGGAAATAGTTG 451
OY      61 HisPheAspIleLysProProLeuHisHisAlaLysAsnGlyAspPheSerPheThrGly 80
DB      450 CATTTTGATTTCAAACTCCCACTGATGATTAATAAATGATTTTCTTTTCTGGA 391
OY      81 LeuGlnHisValThrAspLysIleIleMetLysGlnLysGlnGlyIleGluLys 100
DB      390 CTTCAACATGATTATTAATTAAGACATTAATGCAAGAGAAAGAGAGGATTATGAGAG 331
OY      101 GlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisThrMetAlaCys 120
DB      330 GGGCAAAATCTGCTTCACTGACAGCATTTGCTGCTGACAGTACAGACAGAGAGCTTGC 271
OY      121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuPro 140
DB      270 CATATTGCAGAAAAGAACTCATCTGCTATTTCTTTGCAAAACAGAGAGACTTATTATGT 211
OY      141 GlnAsnAsnAlaValLeuValAlaSerGlyGlyValAlaSerAsnPheTyrIleArgArg 160
DB      210 CAAAGTAATGCAATATGTTGATCTGAGAGGTTGCAAGTAATTAATTAACGAAA 151
OY      161 AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProArgLeu 180
DB      150 GCTTTGAAGATTGTGCAATGCAACAAATGCACTCTGTTGTCTCTCCACAGCTA 91
OY      181 CysThrAspAsnGlyIleMetIleAlaTyr 190
DB      90 TGCACTGATTAATGCAATTAATGATTCATGG 61

```

RESULT 9

CK941819/c CK941819 637 bp mRNA linear EST 15-MAR-2004
LOCUS 4065407 BARC 10BOV Bos taurus cDNA clone 10BOV12_F24 3', mRNA
DEFINITION sequence.
ACCESSION CK941819
VERSION CK941819.1 GI:45456199
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
1 (bases 1 to 637)
REFERENCE Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Hartney,
G.P., Bosak, S., Rubenfield, M. and Gabarrere, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut
Unpublished (2004)
JOURNAL Contact: Tad S. Sonstegard
Bovine Functional Genomics Laboratory
Animal and Natural Resources Institute
Bldg. 200 Km2A BARC-Bast, Beltsville, MD 20705, USA
Tel: 3015048416
Fax: 3015048414
Email: tads@anri.barc.usda.gov

Single pass sequencing. Bases called and trimmed with phred
0.000925 using options -trim alt -trim fast. Vector identified
by cross_match using options -mismatch 12 -minscore 18
Plate: 12 row: F Column: 24
Seq primer: AGCGATACACATTCACACAGG
High quality sequence stop: 637.
Location/Qualifiers

FEATURES

source
1..637
/organism="Bos taurus"
/mol_type="mRNA"
/strain="Holstein"
/db_xref="taxon:9913"
/clone="10BOV12_F24"
/sex="Male"
/tissue_type="Pooled"
/dev_stage="Multiple"
/lab_host="DH10B T1 Phage resistant"
/clone_lib="BARC 10BOV"
/note="Organ: Small Intestine; Vector: pagen-1; Site_1:
EcoRV; Site_2: NotI; Equimolar amounts of mRNA extracted
from proximal jejunums of 18 and 21 wk old steers, and
distal ileums of 14 day old calves. proximal jejunum
exposed to C. oncophora for 3 and 6 weeks, and distal
ileum exposed to C. parvum for 7 days"

ORIGIN

Alignment Scores:
Pred. No.: 7.81e-87 Length: 637
Score: 838.00 Matches: 153
Percent Similarity: 92.47% Conservative: 19
Best Local Similarity: 82.26% Mismatches: 14
Query Match: 67.58% Indels: 0
DB: 7 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x CK941819 (1-637)

```

OY      54 AlaLysGlnGlyAsnArgPheHisPheAspIleLysProProLeuHisHisAlaLysAsn 73
DB      632 GCCAAACAGGAAATTAATATTCATTTGATTTCCAGCTCCCAACAGTCAAAAT 573
OY      74 CysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysGln 93
DB      572 TGTGATTTCTTTTCTTGAAGCTTCAACAGCTTAATTAATGATTAATGCAAAAGCA 513
OY      94 LysGlnGlnGlyIleGluLysGlnIleLeuSerSerAlaAlaAspIleAlaAlaThr 113
DB      512 AAGAGAGAGATTCAGAGAGGAGGAGCTCTGCTTCAAGTGCAGACATTTGCTTGGC 453
OY      114 ValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPheCys 133
DB      452 GTGCAGACACACCGGCTCGTCAATTCAGAAAAGAACACATCTGCTCTTCTGTCTGC 393
OY      134 LysGlnArgAspLeuProGlnAsnAsnAlaValLeuValAlaSerGlyGlyValAla 153
DB      392 AAGCAGAGAGGCTTCTTAAGTCAGAGTACGACATCTGTTGATTCGAGGCTGCAC 333
OY      154 SerAsnPheTyrIleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeu 173

```

Db 332 AGTAATTATATATCGAAAAGCCCTGGAATTTGTGACCAATGACACAGTGCCTTTG 273
 QY 174 LeuCyProProProArgLeuCyThrAspAsnGlyIleMetIleAlaTPaanglyIle 193
 Db 272 CTGTGCGCGCCGCCAGACTTTGCACTTGACCAACGGCTTATGATTTGATGGAATGGTGT 213
 QY 194 GUAATGLeuAAlaGlyLeuGlyIleLeuHisAspIleGluGlyIleArgTyrGluPro 213
 Db 212 GAAAGACTACGTGCTGGCTGGCATTTTACACACACAGAGGCAATCCGCTACGAACCA 153
 QY 214 LysCysProLeuGlyValAspIleSerLysGluValGlyGluAlaSerIleLysValPro 233
 Db 152 AAATGTCCTCTTGAGATGATATATATCAAAAAGAGTTGAGAGAGCTGCTATTAAGAATGCCA 93
 QY 234 GlnLeuLysMetGluIle 239
 Db 92 AGATTAAATGAAGATT 75

RESULT 10
 BQ961028 922 bp mRNA linear EST 21-AUG-2002
 LOCUS AGENCOURT_8863711 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6423902
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ961028.1 GI:22376506
 VERSION BQ961028
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 922)
 NIH-MGC http://mgi.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabbs-remail.nih.gov
 Tissue Procurement: DCTD/DTP/Gazdar
 cDNA library Preparation: Rubin Laboratory
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
 Plate: LCM2603 row: d column: 15
 High quality sequence stop: 584.
 Location/Qualifiers
 1..922
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6423902"
 /tissue_type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 18"
 /note="Organ: lung; Vector: pOT7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 4.6e-82 Length: 922
 Score: 799.00 Matches: 179
 Percent Similarity: 88.04% Conservative: 5
 Best Local Similarity: 85.65% Mismatches: 16
 Query Match: 64.44% Indels: 10
 DB: 5 Gaps: 3

US-10-649-273-2_COPY_176_414 (1-239) x BQ961028 (1-922)

QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
 Db 291 CTGTGGCATATTAGTTCAAGAGGTTTCAGATTTTCTGCTTTCGAAATCTTTGGACATA 350
 QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
 Db 351 GCACCAAGGTGACAGCTTGACAGAGTGCAAGAGACTTTCTTTATATTAACATCCAGAG 410
 QY 41 Cys-SerThrMetSerGlyGlyValAlaGlnHisLeuAlaLysGlnGlyAsnArgPhe 60
 Db 411 TGCATCCACCATGAGTGGTGGGAAGCATGAAACATTGGCCAAACAGGAATAGATT 470
 QY 60 eHisPheAspIleLysProProLeuHisHisAlaLysAsnCyAspPheSerPheThrG1 80
 Db 471 TCATTTGACATCAACCTCCCTTCATCATGATAAAATTTGATTTTCTTTTACTCG 530
 QY 80 yLeuGlnHisValThrAspLysIleIleMetLysLeuLysGluGluGlyIleGlyLys 100
 Db 531 ACTTCAACACCTTACTGATTAATAATTAATTAATAAAAGAAAGAGGATTAGAGAA 590
 QY 100 sGlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnHisPheMetAlaCy 120
 Db 591 AGGCAAAATCCTGTTCTTCAAGACAGACATTCCTCCACATGACGACACAAATGGCATG 650
 QY 120 sHisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuPr 140
 Db 651 TCATCTTGTAAGAAAGAA-CATCGGGCTATTCTGTTTGTAAAGACAGACTTGCTGACC 709
 QY 140 oGlnAsnAsnAlaValLeuValAlaSer-GlyGlyValAlaSerAsnPheTyrIleArgA 160
 Db 710 TCAAAATTAATCACTACTGCTGATCTGCGGGGGGTGCAAGTAACTTATATATCCCGC 769
 QY 160 rGlnAlaLeuGlnIleLeuThr-AsnAlaThrGlnCysThr-LeuLeuCyProProArg 179
 Db 770 AAACCTGGAATAATTTACAAACCCACACAGGGGACCTTGTGGTCCCTCCCA 829
 QY 179 gLeuCyThr-AspAsnGlyIleMetIleAlaTTP-----AsnGlyIleGluArg--L 196
 Db 830 ACTATGCACTGATTAATGG---CATTTATGATGCTGCGGAGGAAATGAAAAATA 886
 QY 196 euArgAlaGlyLeuGly 201
 Db 887 CMTGCTCGCTTGGGG 903

FEATURES

source

1..922
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6423902"
 /tissue_type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 18"
 /note="Organ: lung; Vector: pOT7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

RESULT 11
 AV602901/c 545 bp mRNA linear EST 27-NOV-2001
 LOCUS AV602901 Bos taurus kidney fetus Bos taurus cDNA clone EIK1013A07
 DEFINITION 3', mRNA sequence.
 ACCESSION AV602901
 VERSION AV602901.1 GI:9725227
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 545)
 Takasuga, A., Hirotsune, S., Itoh, R., Jitchozono, A., Suzuki, H., Aso, H., and Sugimoto, Y.
 Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs
 Nucleic Acids Res. 29 (22), E108 (2001)
 JOURNAL MEDLINE
 PUBMED 11713328
 COMMENT Contact: Yoshikazu Sugimoto
 Animal Genetics Division
 Shikawa Institute of Animal Genetics
 Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
 Tel: 81-248-25-5641
 Fax: 81-248-25-5725

Email: kazusugi@cocoa.ocn.ne.jp
 Single pass sequencing.
 This clone was obtained from a polyA-deleted cDNA library.
 Location/Qualifiers
 source
 1. 545

/organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /clone="BK1013A07"
 /tissue_type="kidney"
 /dev_stage="fetus"
 /lab_host="DH10B"
 /clone_lib="Bos taurus kidney fetus"
 /note="Vector: pZL1; Site 1: SalI; Site 2: NotI; Poly A was deleted from a NotI site"

ORIGIN

Alignment Scores:

Pred. No.: 3.18e-81 Length: 545
 Score: 789.00 Matches: 147
 Percent Similarity: 90.1% Conservative: 17
 Best Local Similarity: 80.77% Mismatches: 17
 Query Match: 63.63% Indels: 1
 DB: 1 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x AV602901 (1-545)

57 GYAANAATGPhenIshpheaspielysProProleuHishis-AlalysanCysaerph 76
 544 GGAATAATGATGATTTATATTCANCMTCACATGACAGCGTGTAAATATGATTTT 485
 76 eSerPheThnGlyLeuGlnHisValThrAspLysIleIleMetLysGlyGluGluG 96
 484 TTCTTTTGTGACATTCACACCTTATGTATAGATGATATGCAAAAGAAAAAGGA 425
 96 uGlyIleGlyGlyGlnIleLeuSerSerAlaAlaAspIleAlaAlaThrValGlnH 116
 424 AGGATATCGACAGGGGAGGCTCTGTCTTCACGCGGACATGCTGTGCGGACGCA 365
 116 eThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPheCysGlnAr 136
 364 CACCGGGGCTGCCCATTCGCAAAAAGAACACATCGTCTCTGTTCGACAGCAG 305
 136 GAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyValAlaSerAsnph 156
 304 AGGCTTCTTACATCAAGATGACAGGACTGTTGATCTGAGAGGCGTCGAAAGTACTT 245
 156 eTyrlleArgArgAlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysPr 176
 244 ATATATCCGAAAAAGCCCTGGAATTTGACCAATGACACACAGTGCCTTGTGTCGCC 165
 176 oProProAlaGluCysThrAspAsnGlyIleMetIleAlaIleThrAsnGlyIleGluArgLe 196
 184 GCGCCCCCAACACTCTGACATGACACACGCGCTGATGATTCATGAAATGTTGAAAGACT 125
 196 uArgAlaGlyLeuGlyIleLeuHisAspIleGlyGlyIleArgTyGluProLysCysPr 216
 124 AGGTGCTGGCTTGGGCAATTTTACACACACAGAGGATTCGCTACACACCAAAATGTC 65
 216 oLeuGlyValAspIleSerIleGlyValAlaGlyIleAlaSerIleLysValProGlnLeuL 236
 64 TCTTGAGATGATATATCAAAAGAGTTGAGAGAGCTGCTATATAAGTGCACAGATTAA 5
 236 swet 237
 4 AATG 1

RESULT 12
 CB852881
 LOCUS CB852881 490 bp mRNA linear EST 22-Apr-2003
 DEFINITION UI-CF-FNO-afo-b-03-0-UI-s1 UI-CF-FNO Homo sapiens cDNA clone
 ACCESSION UI-CF-FNO-afo-b-03-0-UI 3', mRNA sequence.
 CB852881

VERSION CB852881.1 GI:30047942
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 TITLE
 Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteleostomi; Primates; Catarrhini; Hominiidae; Homo.
 Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: McCray, PB

McCrack Lab

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).

Seq primer: M13 FORWARD

POLYA=NO.

FEATURES
 source
 Location/Qualifiers

1. 490
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-CF-FNO-afo-b-03-0-UI"
 /tissue_type="Human Lung Epithelial cells"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-CF-FNO"
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-FNO is a subcloned cDNA library derived from two normalized Human Lung epithelial cell libraries (EN1 and DU1) The library was subcloned according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bento-soares@uiowa.edu TAG_SEQ=None Found"

ORIGIN

Alignment Scores:

Pred. No.: 1.05e-80 Length: 490
 Score: 784.00 Matches: 148
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 63.23% Indels: 0
 DB: 6 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x CB852881 (1-490)

69 HisHisAlaLysAsnCysAspPheSerPheThrGlyLeuGlnHisValThrAspLysIle 88
 40 CATATGCTATAAATTTGATTTTCTTTTACGACTTCACACGTTATGATTAATA 99
 89 IleMetLysGlyGlyGlyGlyIleGluLysGlyGlnIleLeuSerSerAlaAla 108
 100 ATATGAAAAAGAAAAAGAGGATTCAGAGAGGCAAAATCTGTCTTCAGACACA 159
 109 AspIleAlaIleThrValGlnHisIleThrMetAlaCysHisLeuValLysArgThrHisArg 128
 160 GACATTCCTGACAGACAGACACACATGACATGATCTTGTGAAAAAGACATCGG 219
 129 AlalLeuPheCysLysGlnArgAspLeuProGlnAsnAsnAlaValLeuValAla 148

Db 220 GCTATTCTGTTTGTAGACAGAGACTTGTACTTCAAAAATATGACTGTTGCG 279

Qy 149 SerGIgLVAlAlaserapheryllleargArAlaleuGIuileuthrAsnAla 168

Db 280 TCTGGTGGTGGCAAGTAATCTTATATCCGACAGCTCTGGAANTTTTAAACAAGCA 339

Qy 169 ThrgInCystrhrleuLeuCystrProProArXleuCystrAsphangilyllemetile 188

Db 340 ACACAGTGCATTTGTTGTCTCTCCACAGACTATGACTGATATATGCAATATATGATT 399

Qy 189 AlAtPpAsngilyllegluArgluAlaAglyleuGIuileuHiseapillegluGIy 208

Db 400 GCATGGAAAGTAAATTAAGAAGTACGCTGCTGGCAATTTTACATGACATAGAAGGC 459

Qy 209 lIeArGIyGIuProluCystrPro 216

Db 460 ATCCGCTATGAACCAAAATGTCCTC 483

RESULT 13

LOCUS CF257246 822 bp mRNA linear EST 07-AUG-2003

DEFINITION phn008_902 PHA-activated splenocytes Gallus gallus CDNA, mRNA

ACCESSION CF257246

VERSION CF257246.1 GI:33490501

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 822)

Witzell,H., Bed'Hom,B., Morin,V., Young,J.R., Whitaker,C.J., Chausse,A.M. and Zoorob,R. A collection of chicken ESTs from activated immune cells Unpublished (2003)

TITLE JOURNAL

COMMENT Contact: Zoorob R UPR 1983

FEATURES

source

1..822

Location/Qualifiers

/organism="Gallus gallus"

/mol_type="mRNA"

/db_xref="taxon:9031"

/cell_type="splenocytes"

/clone_id="PHA-activated splenocytes"

/note="Vector: pTribLEX2"

ORIGIN

Alignment Scores:

Pred. No.: 2.85e-80 Length: 822

Score: 783.00 Matches: 150

Percent Similarity: 79.57% Conservative: 37

Best Local Similarity: 63.83% Mismatches: 48

Query Match: 63.15% Indels: 1

DB: 7 Gaps: 0

US-10-649-273-2_COPY_176_414 (1-239) x CF257246 (1-822)

Qy 1 leuLeuAlaleuAlaInglyValSerAspPheLeuLeuGIyLysSerLeuAspIle 20

Db 119 ATCTTGCACTAGACACAGAGATTTCAGATTTCCTTCTGCTGACACTCATATGATACA 178

Qy 21 AlaProGIyAspPheLeuAspLysValAlaArgArgLeuSerLeuIleLysHiseProGIu 40

Db 179 GUCUCAGTACATGTTGGATAGTAGTACAGAAAGCGCTCTTTAGTGAACACCCGAG 238

Qy 41 CySerThrMetSerGIyLysAlaIleGIuHiseuAlaLysGIyLysAsnArgPhe 60

Db 239 TGCCACGACATGCGCGGGGGAGAGCAATAGACACCTGGCTCAAAACCGAGACTGGCA 298

Qy 61 HisePheAspIleLysProProLeuHiseAlaLysaEncyAspPheSerPheThrcly 80

Db 299 CAGTACACTTTCACACTTCCCAACAGATGATGATCTGATTTTCTTCTCCCGCA 358

Qy 81 LeuGIuHiseValThrAspLysIleIleMetLysLysGIuLysGIuGIuGIuLys 100

Db 359 CTTCAAGCGCTTTCACAAAGCCATCTTCAAGAAAGAAAGAAAGAAAGATTCACAA 418

Qy 101 GlyGIuHiseuSerSerAlaAlaAspIleAlaIleThrValGlnHiseThreAlaCyS 120

Db 419 GGGGAATTCCTGCTTCGCTTAAGACATCCCTGCTGTCGACACAGACTAGTGTCTCT 478

Qy 121 HiseValLysArgThrHiseArgAlaIleLeuPheCyLysGlnArgAspLeuPro 140

Db 479 CATATTTCAGCGGACACACCGACCATGCTCTTCTGCAAGAAACAGCATATATTA 538

Qy 141 GlnAsnAsnAlaValLeuValAlaSerGIyLysValAlaSerAspPheTyrlleArgArg 160

Db 539 CCAAAACTGCACCTGCTGTTGTATCAGAGAGAGTGCAGATTAATCAGTATATCAGAAA 598

Qy 161 AlaleuGIuileuThrAsnAlaThrcInCystrhrleuLeuCystrProProArgArg 180

Db 599 GCACTGCAGACTCTGGCAAAATGCAACCGTTTCTTCTGCTCTCTCCCAAGGCTG 658

Qy 181 CyThrAspAsngilyllemetileAlaIleArgAsngilyllegluArgLeuAlaGIyLeu 200

Db 659 TGCAACCAATATGAGTATATATTCAGATGATGATGATGATGATGATGATGATGAT 718

Qy 201 GIyIleLeuHiseapillegluGIyIleArgTyrgIuProluCystrProleuGIyValAsp 220

Db 719 GGTATTTTATACAGTACATGATGATGATGATGATGATGATGATGATGATGATGAT 778

Qy 221 lIeSerLysGIuValGIyGluAlaSerIleLysValProGlnLeu 235

Db 779 ATTCCAAAGAGTTGAA-GAGGATTCATCAAGTCCCAAGTCA 822

RESULT 14

LOCUS BUI27463 866 bp mRNA linear .EST 25-NOV-2002

DEFINITION 603114407F1 CSECHU20 Gallus gallus CDNA clone CHEST6m2 5', mRNA

ACCESSION BUI27463

VERSION BUI27463.1 GI:25338728

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 866)

Boardman,P.E., Sanz-Ezquerro,J., Overton,I.W., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J. A Comprehensive Collection of Chicken cDNAs Curr. Biol. 12 (22), 1965-1969 (2002)

TITLE JOURNAL

COMMENT Contact: Simon Hubbard Department of Biomolecular Sciences University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

1..866

Location/Qualifiers

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn, HiseX"

/db_xref="taxon:9031"

/clone="CHEST6m2"

```

/dev stage="j36"
/lab_host="DH108"
/clone_lib="cseqch20"
/notes="Organ: limbs; Vector: pBluescript II KS(+); Site 1:
EcoRI; Site 2: NotI; Modification of pBluescript II KS(+)
[is]tagene] vector to accommodate cDNA produced with the
T-primed protocol (construction of uni-directionally
cloned cDNA libraries from messenger RNA for improved 3'
end DNA sequencing by Glenn Pl, et al. U.S. Patent #
6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI.
Ligate in double stranded adaptor containing BspI and
BamHI sites [5'ggccggcgccagcccgccgagtcgcaaaaaag]
[5'aattctttttccgagtcgggctgcagc]"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 9,58e-77 Length: 866
Score: 753.00 Matches: 149
Percent Similarity: 78.81% Conservative: 37
Best Local Similarity: 63.14% Mismatches: 50
Query Match: 60.73% Indels: 2
DB: 5 Gaps: 0

```

US-10-649-273-2_COPY_176_414 (1-239) x BUI27463 (1-866)

```

QY 1 LeuLeuAlaLeuValGlnGlyValSerAspPheLeuLeuGlyLysSerLeuAspIle 20
DB 25 ATCTGGCAGTAGACAGAGAGATTTCAGATTTCTTCTGCGACAGCTCATATGATA 84
QY 21 AlaProGlyAspMetLeuAspLysValAlaArgArgLeuSerLeuIleLysHisProGlu 40
DB 85 GCACCGAGGTGACATGCTGATAGTAGCAAGAGGCTCTTATGAAAGCACCCGAG 144
QY 41 CysSerThrMetSerGlyLysAlaIleGlnHisLeuAlaLysGlnGlyAsnArgPhe 60
DB 145 TCCACAGGATGGCTGGGGGAAAGCAATAGACAGCTGCTCAACCGAGAGCTGGCAA 204
QY 61 HisPheAspIleLysProPheLysHisAlaLysAsnGlyAspPheSerPheThrGly 80
DB 205 CAGTACACTTTCAGATTTCCCATGCAACAGTATGTAAGTATTTCTTCTCCGCA 264
QY 81 LeuGlnHisValThrAspLysIleIleMetLysGlyLysGlnGlyIleGluLys 100
DB 265 CTTCGAGAGCTTGTCAACAAAGCCATTTCTTCAAAAGAAAGAGAGATTCAGAA 324
QY 101 GlyGlnIleLeuSerSerAlaAlaAspIleAlaIleThrValGlnHisThrMetAlaCys 120
DB 325 GGGGAAATCTGCTGCGCTTAAGACATGCTGCTGCACAGCATGTAGTGGCTGCT 384
QY 121 HisLeuValLysArgThrHisArgAlaIleLeuPheCysLysGlnArgAspLeuLeuPro 140
DB 385 CATATATTCACAGCGGACACACAGCAGCATCTTCTGCAATGAAAAACACATATTTTA 444
QY 141 GlnAsnAlaValLeuValAlaSerGlyValAlaLysSerAsnPheThrIleArgArg 160
DB 445 CCAAAAACAGCAACTGCTGTTGATCAGAGAGAGTTCAGATATCATCATCAGAAA 504
QY 161 AlaLeuGlnIleLeuThrAsnAlaThrGlnCysThrLeuLeuCysProProPheArgLeu 180
DB 505 GGACTGCAAGACTCTGGC-AATGCAACGGTTTGTCTTCTGCTCTCCCAAGGCTG 563
QY 181 CysThrAspAsnGlyIleMetIleAlaTPAsnGlyIleGlnArgLeuAArgAlaGlyLeu 200
DB 564 TCACCGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 623
QY 201 GlyIleLeuHisAspIleGlnGlyIleArgGlyArgProLysCysProLeuGlyValAsp 220
DB 624 GGTATATTA-TACAGTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 682
QY 221 HisSerLysGlnValGlyLysAlaSerIleLysValProGlnLeuLys 236
DB 683 ATTTCCAAAAGAGTTGAAGAGATTCATCAGAGTCCAGACTTAAG 730

```

RESULT 15

CK365185

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

EST

Rattus norvegicus (Norway rat)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Scurionathii; Muridae; Murinae;

Rattus.

1 (bases 1 to 736)

AUTHORS

TITLE

JOURNAL

COMMENT

NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-remail.nih.gov
Tissue Procurement: Howard Jacobs
CDNA Library Preparation: Express Genomics
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LNL4944 Row: d Column: 02
High quality sequence stop: 736.
Location/Qualifiers

FEATURES

source

1..736

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="IMAGE:7097620"

/tissue_type="lung, pooled"

/lab_host="DH10B Tona"

/clone_lib="NIH_MGC_231"

/note="Organ: lung; Vector: pExpress-1; Site 1: EcoRV;

Site 2: NotI; RNA obtained from pooled lung tissue from a

mix of male and female animals at 8 wk old. Tissues were

snap-frozen and kept at -80C for two days before RNA

extraction and purification (TRI-reagent method). cDNA was

primed using oligo-dT primer:

5'-pGACTTACTTTCAGTTCGACGCGCCGCTT-25-3' and cloned into

the EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb

resulted in an average insert size of 2.3 kb. This primary

library is not normalized (normalized primary library is

NIH-MGC 232) and was constructed by Express Genomics

(Frederick, MD). Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:

```

Pred. No.: 2.22e-76 Length: 736
Score: 749.00 Matches: 141
Percent Similarity: 91.02% Conservative: 11
Best Local Similarity: 84.43% Mismatches: 15
Query Match: 60.40% Indels: 0
DB: 7 Gaps: 0

```

US-10-649-273-2_COPY_176_414 (1-239) x CK365185 (1-736)

```

QY 73 AsnGlyAspPheSerPheThrGlyLeuGlnHisValThrAspLysIleIleMetLysLys 92
DB 1 AACTGATATTTCTTTTACGGGACTTCAACATGCTCACCGTATGCTATACACACAG 60
QY 93 GlyLysGlnGlyIleGlyLysGlyGlnIleLeuSerSerAlaAlaAspIleAla 112
DB 61 GAAAGAAAGAGCATTTGAGAGGGGCAATCTGTCATCAGCGCGAGACATGCTGCT 120
QY 113 ThrValGlnHisThrMetAlaCysHisLeuValLysArgThrHisArgAlaIleLeuPhe 132

```

```
Db      121 GCGGTACAGCAGCCAAACAGCGGTGCGACCTTGCGAAAAGAACACATCGTGCTATTCTGTT 180
Qy      133 CysLeuGlnArgAspLeuLeuProGlnAsnAsnAlaValLeuValAlaSerGlyVal 152
Db      181 TGCAGCAGAAAATTGGCTATCTCCAGCTAACGCCAGTATTAGTTGTGTGAGGTGTT 240
Qy      153 AlaSerAsnPhenylleArgArgAlaLeuGlnleuThrAsnAlaThrGlnCysThr 172
Db      241 GCAAGTACTTGTACATCCGAGAGCATTTGGAATTTGACCAATGCAACACATGCAC 300
Qy      173 LeuLeuCyseProProProAlaArgLeuCyseThrAspAsnGlylleMetlleAlaTrpAsnGly 192
Db      301 TTGTTGTGTCCTCCGAGACTGTGCATGACATATGATCATGATTCATGGAATGGA 360
Qy      193 lleGlnArgLeuArgAlaGlyLeuGlylleuHisAspIleGlnGlylleArgTyrglu 212
Db      361 ATTGAAGATTACGTGCTGGCTTGGCATTTTACATGATGAGAAGACATCGATACGAA 420
Qy      213 ProLysCyseProLeuGlyValAspIleSerLysGluValGlyGluAlaSerIleLysVal 232
Db      421 CCAAAATGTCCTCTCGAATATAGACATATCCAGAGAAAGTTGCAGAAAGCTGCCATAAAGTA 480
Qy      233 ProGlnLeuLysMetGluile 239
Db      481 CCAAGATTAAATGACACTT 501
```

Search completed: June 17, 2005, 01:30:06
Job time : 2054.13 secs